THE COGNITION OF MULTIAIRCRAFT CONTROL (MAC): COGNITIVE ABILITY PREDICTORS, WORKING MEMORY, INTERFERENCE, AND ATTENTION CONTROL IN RADIO COMMUNICATION

Kelly M. Amaddio, Second Lieutenant, USAF

AFIT-ENV-MS-15-M-205

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

THE COGNITION OF MULTIAIRCRAFT CONTROL (MAC): COGNITIVE ABILITY PREDICTORS, WORKING MEMORY, INTERFERENCE, AND ATTENTION CONTROL IN RADIO COMMUNICATION

THESIS

Presented to the Faculty

Department of Systems Engineering and Management

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Engineering Management

Kelly M. Amaddio

Second Lieutenant, USAF

March 2015

AFIT-ENV-MS-15-M-205

THE COGNITION OF MULTIAIRCRAFT CONTROL (MAC): COGNITIVE ABILITY PREDICTORS, WORKING MEMORY, INTERFERENCE, AND ATTENTION CONTROL IN RADIO COMMUNICATION

Kelly M. Amaddio
Second Lieutenant, USAF

Committee Membership:

Dr. Michael Miller
Chair

Dr. John Elshaw
Member

Dr. Victor Finomore
Member

Abstract

As the number of U.S. Air Force missions requiring UAVs has rapidly increased without commensurate increases in manpower, systems which permit a single operator to supervise and control multiple, highly-automated aircraft are being considered. The operator of such a system may be required to monitor and respond to voice communications for multiple UAVs, each of which can have aircraft specific call signs. The need to monitor this array of call signs may impose excessive requirements on constrained operator attention, working memory, and cognitive processing. The current research investigates the cognitive load (number of aircraft call signs) an individual can handle and explores the effect of proactive interference (PI) within this application. The results indicate a reduction in performance as the number of call signs are increased from 5 to 7 in the presence of PI. Additionally, this study seeks to understand if individual differences in working memory and attention predict performance on the multiaircraft control radio communication task through the application of the Operations Word Span test, Attention Control Scale, and GRE scores. Hierarchical linear modeling was used to determine the relationships among these and other variables.

Acknowledgments

First of all, I want to thank God for being with me through all of this. Thank you to my husband and my parents for their continuous support. Thank you to my lab mates Danielle, Stephen, and Dave for being constantly uplifting. A special shout out to Cheeto Puffs for being the highlight of my day on many days! Thanks to all of my classmates who have been amazing mentors and friends throughout my second lieutenant years; I have learned so much from you guys. Thank you to my advisor, Dr. Miller for providing so much guidance and support throughout this process and to my committee members Dr. Elshaw and Dr. Finomore for their feedback and help. I couldn't have done it without all of you!

Kelly M. Amaddio

Table of Contents

List of Figures

List of Tables

THE COGNITION OF MULTIAIRCRAFT CONTROL: WORKING MEMORY, PROACTIVE INTERFERENCE, AND ATTENTION CONTROL IN RADIO COMMUNICATION

I. Introduction

General Issue

The United States military is currently involved in many conflicts and humanitarian relief efforts worldwide. As these activities continue and budget pressures force reductions in the number of military personnel, technology is increasingly applied as a force multiplier (Unmanned Systems Integrated Roadmap FY2013-2038). Unmanned Aerial Vehicles (UAVs) have become increasingly important in recent years as they significantly enhance the gathering of intelligence, surveillance and reconnaissance (ISR) without risking bodily injury to the operators. Additionally, UAVs have been proven useful in a multitude of civilian contexts, including meteorology, wildlife preservation, agriculture, search and rescue, and logistics (Handwerk, 2013). However, the number of UAV sorties has increased exponentially in recent years despite the limited number of pilots available to control them.

As a result, new military concepts of operation (CONOPS) are under consideration wherein a single pilot might control multiple aircraft during certain phases of flight. Currently, one operator at the airbase handles all ground operations, launching, and retrieving all UAVs, as shown in Figure 1. However, once the aircraft is out of the line of sight, another operator takes control of the vehicle, completes the mission, and then returns it to the ground station controller.

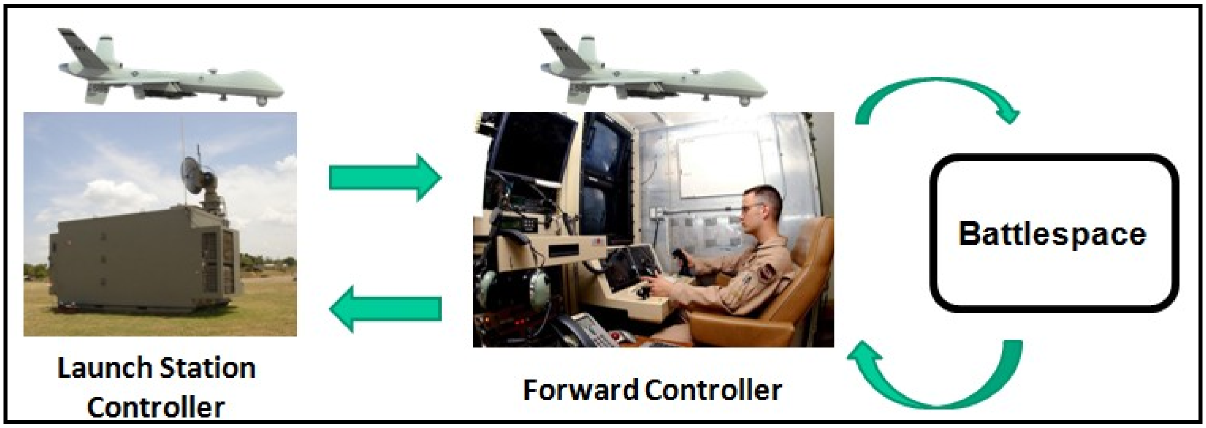

Figure 1. Old UAV CONOPS.

A proposed change to this model (Figure 2) is to employ transit operators to simultaneously pilot multiple, semi-autonomous aircraft between the airbase and the battlespace. This model would appear to be beneficial because the transit of aircraft to and from the battlespace is relatively low in task load, so the operator would have the cognitive resources to pilot multiple UAVs simultaneously (Columbi et al., 2012). Since a single UAV sortie could last for several hours and the transit from the launch station to the battlespace could comprise a significant proportion of that time, it is important to consider adding an operator to control multiple aircraft during this time.

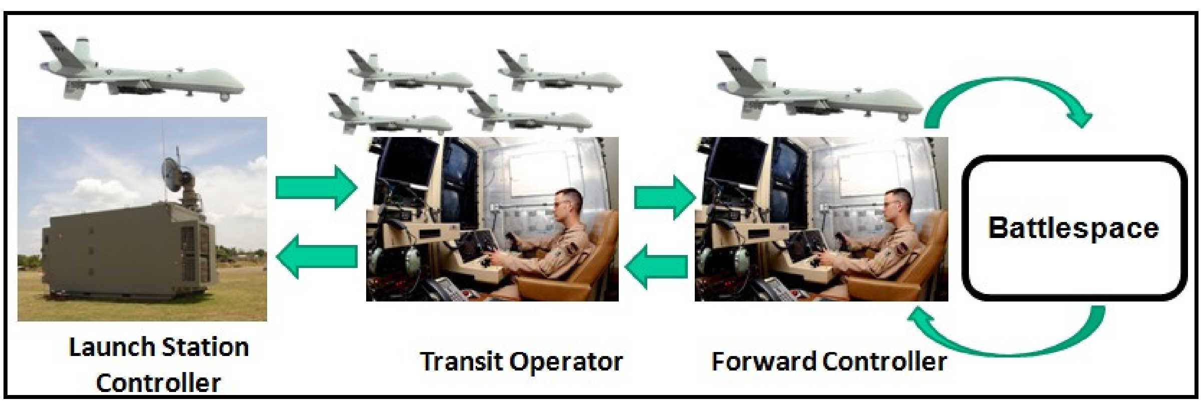

Figure 2. Multiaircraft Control (MAC).

MAC would also result in manpower savings, although diminished incremental savings is predicted to occur as aircraft to pilot ratio increases (McGrogan, et al., 2011).

There are a few mission considerations when employing the multiaircraft control model. For the pilots to operate multiple aircraft at once, they will have to monitor and respond to the call signs and radio communications of each aircraft, potentially requiring a significant increase in radio communications for the transit operator as compared to today's operators. Additionally, there is a concern that proactive interference may occur when pilots transfer aircraft to other pilots, but still hear the previous aircraft specific radio calls. In this environment, the operators must actively ignore irrelevant radio communications while responding to only the call signs they are responsible for at a given moment in time. Several principles related to working memory, interference, and attention are important in the analysis of this issue.

Finally, selecting individuals to perform UAV operations is important when operating within constrained budgets. By using several measures of cognitive ability, which are predictive of an operator's performance on a multiaircraft control task, training costs can be reduced by selecting only operators who are likely to succeed before they enter the career field.

Problem Statement

The current military mission requirements dictate the need for multiaircraft operations. This is a new concept of operation, so it is important to explore the cognitive aspects of attending to multiple aircraft. Additionally, individual differences in working memory capacity, attention control, and cognitive ability could be used to guide personnel selection. While many laboratory experiments have been conducted on working memory, attention, and proactive interference, this research has not been applied to a multiaircraft control scenario where it is unclear as to whether they are related since this task is not similar to traditional working memory tasks.

Research Objective

The objective of this research is to show the cognitive limitations and effects of multiaircraft control, especially in the realm of radio communication for transit operations. This objective will be accomplished by conducting a human subjects laboratory experiment to explore the number of call signs an operator can actively monitor and the impact of proactive interference. Measures of attention control, working memory, and cognitive ability will be analyzed in relation to these results to find predictors of performance.

Research Focus

This research is focused on the cognitive aspects of multiaircraft control including working memory, attention, proactive interference, and cognitive load. Because of this, the performance task and measurements will be more simplistic than real-world UAV operations to isolate the cognitive mechanisms involved in these processes.

Investigative Questions

The successful completion of the research objective will occur when the following questions are answered:

1. How is performance on a radio communications task related to cognitive demand (number of call signs and presence of proactive interference)?
2. Is an individual's level of attention control (attention focusing/attention shifting) related to their performance on a radio communications task in which the operator is monitoring multiple call signs?
3. Is an individual's working memory capacity related to their performance on a radio communications task involving multiple call signs?
4. What aspects of an individual are the most predictive of performance on a radio communications task: working memory capacity, attention control, GRE scores?

The causal diagram in Figure 3 depicts the predicted relationships between the variables of interest. Since the cognitive measures are supposed to measure the individual cognitive traits, a positive relationship is expected between them where the Attention Control Scale measures attention control, the Operations Word Span Test measures working memory capacity, and GRE scores measure cognitive ability (although these relationships will not be directly tested in the current study). These cognitive measures are also expected to be positively related to one another and positively related to performance. As experimental conditions are manipulated (addition of call signs or presence of proactive interference), the individual's perceived workload and difficulty levels are expected to increase and this is expected to be negatively related to performance.

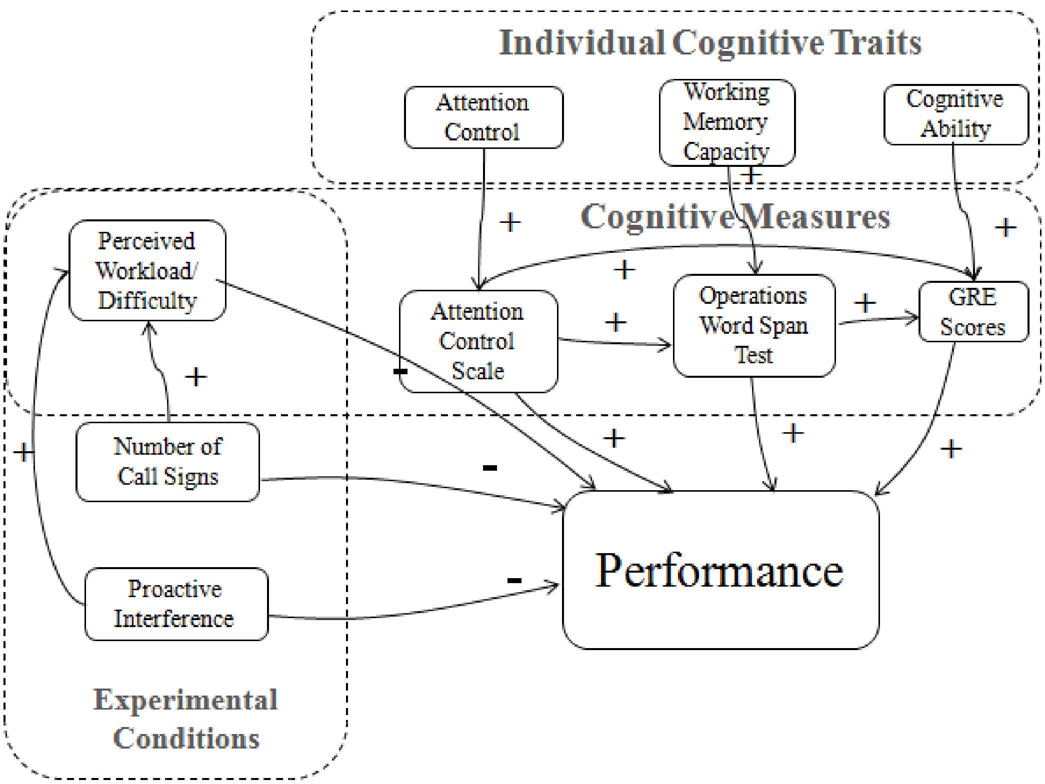

Figure 3. Predicted causal diagram.

Methodology

This study will employ a four-trial human subjects mulitaircraft control laboratory experiment using the Air Force Research Laboratory's Multimodal Communication Suite (MMC), the Operations Word Span Test, Attention Control Scale, GRE scores, and a post-trial questionnaire to answer the research questions. Throughout the experiment, accuracy scores, response times, workload, and subjective difficulty scores will be collected.

Assumptions and Limitations

This experiment thoroughly investigated the cognitive aspects of multiaircraft control; however, there were a few assumptions and limitations to consider. Due to time and test personnel constraints, the number of available participants were limited, resulting in a small sample size. In the design of this experiment, assumptions were made regarding the operational components of the UAV control task. The workload level was assumed to be high enough where the participants had to intentionally process the information, but not so high that they could not hear all of the information. Therefore, radio calls were played at a consistent rate. This differs from the operational environment, which would contain variable levels of workload and radio calls occurring at different speeds. The MMC software interface used during the experiment includes auditory and visual components that are not reminiscent of the operator interface in actual UAV operations. It is assumed, however, that the results of this laboratory situation can be extrapolated to the operational environment. Further, the experimental paradigm did not include any secondary tasks, despite the fact that the pilots in an operational environment would be responsible for other tasks like navigation, communication, and aircraft monitoring. This simplification of the environment, however, made it possible to assess the ability of the operators to perform this auditory task under near ideal circumstances.

Implications

Many studies have been conducted to explore the effect of cognitive load on task performance. Few have been conducted in a UAV environment and none address the proactive interference that may occur as a result of changing attention from one set of aircraft call signs to another. The results of this study will inform system designers and policy makers regarding the level of cognitive demand an individual can handle given personal characteristics and technology limitations. Additionally, the relationship between performance on this multiaircraft and scores on other cognitive measures could be used to select personnel who would be successful at this task.

Preview

This chapter addressed the objectives and focus of this research, an overview of the method, assumptions and limitations, and implications of the study. Chapter 2 is an article that was accepted and submitted to the 2015 International Association of Aviation Psychology (ASAP) conference, which contains results from the MMC task, including accuracy scores, response times, and effects of proactive interference. Chapter 3 contains a draft journal article to be submitted to the International Journal of Aviation Psychology, which explores the relationship between several cognitive tests (Attention Control Scale, Operations Word Span Test, GRE scores) and performance on a simulated multiaircraft control task. Finally, Chapter 5 contains a summary of this research study, overall conclusions, and recommendations for future research.

II. THE COGNITION OF MULTI-AIRCRAFT CONTROL (MAC): PROACTIVE INTERFERENCE AND WORKING MEMORY CAPACITY

Kelly Amaddio, Michael Miller, Ph.D., and John Elshaw, Ph.D.

Air Force Institute of Technology, Dayton, Ohio

Victor Finomore, Ph.D.

United States Air Force Academy, Colorado Springs, Colorado

As the number of U.S. Air Force missions requiring UAVs has rapidly increased without commensurate increases in manpower, systems which permit a single operator to supervise and control multiple, highly-automated aircraft are being considered. The operator of such a system may be required to monitor and respond to voice communications for multiple UAVs, each of which can have aircraft specific call signs, which may impose excessive requirements on constrained operator attention, working memory, and cognitive processing. The current research investigates the cognitive load (number of aircraft call signs) an individual can handle and explores the effect of proactive interference (PI) within this application. The results indicate a reduction in performance as the number of call signs are increased from 5 to 7 in the presence of PI. Interestingly performance with 5 call signs without PI is lower than performance with 5 call signs in the presence of PI.

Introduction

The United States military is currently involved in many conflicts and activities worldwide. As these wars continue and budget pressures forces the decrease of military personnel, technology is relied upon as a force multiplier. Unmanned Aerial Vehicles (UAV) have become increasingly important in recent years as they significantly enhance the gathering of Intelligence, Surveillance and Reconnaissance (ISR) without risking bodily injury to the operators. As a result, the number of UAV sorties has increased exponentially in recent years despite the limited number of pilots available to control them. As a result, new concepts of operation are under consideration wherein a single pilot might control multiple aircraft during certain phases of flight. For example, transit operators may be employed to simultaneously pilot multiple semi-autonomous aircraft between an airbase and the battlespace. If pilots are going to be operating multiple aircraft at once, they will have to monitor and respond to a large throughput of radio communications. Additionally, there is a concern that proactive interference (PI), when previously stored information prevents the learning of new information, may occur when pilots transfer aircraft to other pilots, but still hear the previous aircraft specific radio calls. Several principles related to working memory, interference, and attention are important to the analysis of this issue. The following study is a cognitive laboratory experiment aimed at evaluating cognitive load and the effects of PI.

The ability of an operator to listen to and respond appropriately to radio traffic which contains references to the call signs of the aircraft they are controlling, as well as other entities, is likely to be constrained by their available working memory. Working memory is involved in storing and manipulating information for short-term use in tasks like reasoning and comprehension (Baddeley & Hitch, 1974). A common model of working memory that has been proposed by Baddeley (2000) contains a set of subsystems, including the central executive, which controls attention between the visuospatial sketchpad, episodic buffer, and phonological loop subsystems. The visuospatial sketchpad manipulates visual images while the phonological loop is responsible for storing and replaying words and sounds. The episodic buffer temporarily stores and integrates multimodal information and relays information between the visuospatial sketchpad and phonological loop. The auditory component of this model is important to the current study because participants are asked to listen and respond to a select series of aircraft radio calls.

Although significant research has been conducted on visual working memory, auditory working memory has garnered less attention. Considering this, Kumar et al., 2013 attempted to test auditory working memory over a continuous scale by using sequences of tones in different lengths where participants were asked to adjust a dial to replicate a specific tone that they heard. The findings indicate that increasing the number of tones held in working memory reduced the precision of the memories, much like what is found in visual working memory (Alvarez & Cavanagh, 2004).

Working memory is usually measured by span tasks that require the individual to simultaneously process and remember verbal information, usually words, letters, or numbers. The current study uses a more functional measurement of working memory by requiring the individual to remember a set of words and respond to them when they are spoken in the form of radio calls. They also have to perform this task in the presence of distracting, and sometimes interfering information. This increases their cognitive load, which is considered a measure of the mental effort used to maintain information in working memory (Sweller, 1988), implying that working memory is limited by the amount of information it can hold and process. Miller's (1959) article provides the rule of thumb for information processing capacity: people's ability to process and remember limits them to 7± 2 items. Although the current study only requires participants to recognize call signs (instead of recalling them), the temporal complexity of the task and presence of distracting information causes us to hypothesize that individuals will be able to effectively attend to a similar number of call signs.

One of the primary functions of working memory is to navigate the effects of PI (Kane & Engle, 2000) where timely information replaces less recent information to reduce the likelihood of confusion. Therefore, effective working memory will suppress memory of outdated information to prevent it from interfering with the encoding of new information. PI has been shown to affect performance on working memory tasks. May, Hasher, and Kane (1999) found that performance on a working memory span test was improved when measures were taken to prevent PI (e.g., temporally separating trials). Kane and Engle (2000) found that individuals with low working memory spans showed greater susceptibility to PI under low cognitive load conditions, but under high cognitive load conditions, both high and low working memory span individuals showed equal levels of PI. Engle and Oransky (1999) propose that controlled attention is the mechanism by which working memory functions. They describe controlled attention as "an ability to effectively maintain stimulus, goal, or context information in an active, easily accessible state in the face of interference, to effectively inhibit goal-irrelevant stimuli or responses, or both" (Kane, Bleckley, Conway, & Engle, 2001, p.18). Neurological evidence shows that different information (sensory, semantic, etc.) is stored in different areas of the brain (Postle, 2006) suggesting that working memory should be seen as directing attention towards different memory codes stored in long term memory. Although these models of working memory have different implications for the design of interfaces to support MAC, they all support the view that the operator's attention must be divided between the visuospatial tasks necessary to control the aircraft, processing of audio call signs, and the integration of this information.

The current literature has shown that while working memory tests have been applied in numerous laboratory environments, they have not been applied to understand individual differences in real-life applications of working memory. This study will provide a more functional test of working memory by measuring participants' performance (in terms of accuracy and response time (RT)) on a multiaircraft control task in the presence of distracting information. It is predicted that higher cognitive load (created by the addition of more call signs and the presence of PI) will decrease performance.

Method

Participants

Twenty one (5 female and 16 male) volunteers with ages between 22 and 44 (*M* = 27.75, *SD* = 4.96) participated in the study. Participants were required to have a visual acuity of 20/30 or better, determined using a Logarithmic Near Visual Acuity Chart ("New ETDRS" Charts, 2011) and normal color vision, determined using isochromatic plates(Ishihara, 1980). There was no educational requirement, although most participants were graduate engineering students. Participants were recruited through e-mail. A participant number was assigned to each consenting participant's data and no personally identifiable information was retained per Institutional Review Board Protocol.

Apparatus

The experiment was conducted in a 6ft x 6ft cubicle in a quiet laboratory to minimize distractions. The experimental setup consisted of Bose AE2w headphones and a laptop to present the call signs using the Multi-modal Communication (MMC) software (Finomore, Popik, Castle, & Dallman, 2010). Participants were also given a wireless ten-digit number keypad, a clipboard containing a number grid with four rows and three columns, and a clipboard containing the list of call signs. The list of call signs was provided to the participants to remember before the experiment began and attached to the left wall to the cubicle slightly above eye level once the participants indicated their comfort with the call signs. The placement was selected to require the participant to actively turn their head to view the list.

The Multi-modal communication program (MMC) is an Air Force Research Laboratory developed multi-modal, network-centric communication management suite developed to aid Command and Control operators in increasing communication intelligibility and reduce mental workload. This tool combines several features designed to improve the performance of the users, including spatial audio, speech transcription, data capturing and playback, chat messages, and automatic keyword highlighting (for full description of the MMC tool see Finomore et al., 2012). Additionally, this tool has been used extensively as a research tool to evaluate a variety of communication effectiveness questions (Blair, Rahill, Finomore, Satterfield, Shaw, & Funke, 2014; Finomore, et al. 2010; Finomore, Stewart, Singh, Raj, & Dallman, 2012; Finomore, Satterfield, Sitz, Castle, Funke, Shaw, & Funke, 2012; Santana, Langhals, Miller & Finomore, 2013). This experiment utilized monaural sound, a chat window to prompt the participant to enter the appropriate code, and the logging function to record the participants' inputs.

Experimental Procedure

In the design of this experiment, a few assumptions were made regarding the operational components of the UAV control task. Each aircraft was assumed to have a unique call sign and individuals having different voices made radio calls for any of the call signs (one voice was not reserved for each call sign) as is typical in current operational environments. It was also assumed that the workload level was high enough where the participants had to intentionally process the radio calls but not so high that they could not listen to all of the radio calls. Therefore, radio calls were made every five seconds. This differs from the operational environment, which would contain variable levels of workload. Additionally, there were no secondary tasks to accomplish while participants were completing the auditory task, despite the fact that the operators in an operational environment will be responsible for other tasks like navigation, communication, and aircraft monitoring. This simplification of the environment made it possible to assess the ability of the operators to perform this auditory task under near ideal circumstances.

Upon arrival, participants were randomly assigned to one of two groups based on their participant number. They were given a quick explanation of the software and task, and then given a three minute practice trial where they were responsible for three call signs. This practice trial was designed to minimize the possibility of a learning effect. Although a hearing test was not administered, participants were encouraged to set the volume of the radio calls to their comfort level during this warm-up period.

Based on their group, participants were asked to attend to either 5 or 7 call signs (out of 13 possible call signs) during each of four 8-minute experimental trials. The trials were counterbalanced to offset a potential learning effect. Participants in Group 1 were assigned five call signs for the first two trials and seven for the second two trials. Participants in Group 2 were assigned seven call signs for the first two trials and five call signs for the second two trials. Each

8-minute trial contained 100 radio calls that were evenly spaced 5 seconds apart. Approximately 50 radio calls were critical and an equal number were distracters. The participants did not know what the ratio was, however. During the second and fourth trials, 20 of the distracters were selected to induce PI as they were among the critical call signs in the previous trial. The order of the radio calls and calls signs was randomized. Table 1 presents the trials and the critical and PI call signs for participant Group 1. The scenarios will be referred to as 5-NP (5 call signs, no PI condition), 5-PI (5 call signs, PI condition), 7-NP (7 call signs, no PI condition), and 7-PI (7 call signs, PI condition).

Table 1.
Call signs experienced by the first participant group during each trial. Call signs which were employed to induce PI during Trials 2 and 4 are shown in Bold-Italics for Trials 1 and 3.

Participant Group	**Trial 1** (5-NP)	**Trial 2** (5-P)	Break (15 minutes)	**Trial 3** (7-NP)	**Trial 4** (7-P)
1	Laker	Laker	Working	Charlie	Charlie
	Hopper	Hopper	memory	Gringo	Gringo
	Arrow	Arrow	capacity test	Laker	Laker
	Charlie	Tiger	followed a	Raptor	Raptor
	Gringo	Eagle	break	Viking	Viking
				Arrow	Thunder
				Tiger	Cobra

The participants were instructed to listen for the commands that contained their call sign. Each radio call began with the word "Ready", which was proceeded by a call sign and a command containing a grid coordinate; for example, "Ready Charlie go to blue one now." The color indicates a column in the grid and the number represents a row in the grid. The grid location would then contain a number. For critical call signs, the participants then found the space on the grid that corresponded with the command, and typed the number from the grid location into the MMC chat window. For example, when the participant heard "Ready Charlie go to blue one now," if the participant was responsible for "Charlie" during that trial (Charlie would be on their list of call signs), they would be expected to find the "blue 1" spot on the grid and type the two digit number in that grid location on the keypad. If the participant heard a call sign that was not on their list, they were instructed to type a zero into the chat window. Also, if for some reason they were not sure whether they were responsible for a specific call sign, they were instructed to type a zero. The randomized numbers on the grid were between 10 and 99. Participants were given as much time as they needed to memorize the call sign list before every trial and were instructed to only look at the list of call signs if they forgot them during the trial. The number of times they looked at the call sign list was recorded by the investigator for every trial.

To keep the participants from habituating to certain experimental conditions (call signs and voices), certain measures were taken. First, the list of critical call signs on the clipboard were shuffled for each trial so that they were not in the same order for sequential trials, making it harder to memorize. All trials contained different orders of radio calls, different call signs, and called for different grid locations. Additionally, a new number grid was used for each trial. Finally, a variety of voices made radio calls for every call sign so that the participant could not ignore or attend to a certain call sign based on the speaker. During the experiment, the participant could hear up to 12 different individual's voices and up to 13 different call signs.

Performance Measures

Data was collected during all trials using the logging function in MMC. After each trial, participants were asked to respond to two 5-point Likert Scale questions: one regarding their workload level (Tattersall & Foord, 1996) and the other regarding the perceived difficulty (1= very easy, 2 = easy, 3 = neutral, 4 = difficult, 5 = very difficult). After the last trial, participants were asked to self report the number of call signs they believed they could reliably monitor.

Numerical responses to the MMC task provided by the participants were evaluated for accuracy and RT. For each trial, the accuracy score was calculated by dividing the number of correct responses by the total number of radio calls and multiplying by 100%. Additionally, a PI accuracy percentage correct score was determined by adding the number of correct responses given for the PI call signs divided by the total number of radio calls expected to induce PI for 5-PI and 7-PI conditions. Finally, the average of the participant's RTs were calculated for each trial as the average of the amount of time lapse between the time when the radio call was spoken and the time the participant pressed enter after typing their numerical response. This score did not account for RTs for correct and incorrect responses.

Results

A two-factor repeated measures ANOVA revealed that there was a significant main effect of the number of call signs as well as the interaction between the number of call signs and the presence of PI on accuracy scores on the MMC task ($F(1, 19) = 7.631$, $p = 0.012$), as shown in the left panel of Figure 4. The interaction was further analyzed by applying a single factor repeated measures ANOVA. This analysis revealed that the accuracy scores were significantly different across trials ($F(2.28, 43.31) = 4.307$, $p = 0.016$, partial eta squared = 0.19). Post hoc tests using the Bonferroni correction determined that scores in the 5-PI condition ($M = 97.11\%$, $SD = 3.75\%$) were statistically higher than scores in the 5-NP condition ($M = 93.70\%$, $SD = 3.16\%$) and 7-PI conditions ($M = 91.73\%$, $SD = 6.48\%$). The scores for 5-NP, 7-NP ($M = 94.14\%$, $SD = 6.44\%$), and 7-PI were not significantly different from one another. Therefore, we can conclude that the highest scoring condition occurred when the participants were tasked with 5 call signs in the PI condition. A paired samples t-test indicated that PI accuracy scores were not significantly different between 5-PI (*M* = 95.29%, *SD* = 12.63%) and 7-PI (*M* = 90.25%, *SD* = 15.27%). Additionally, an independent samples t-test showed that accuracy scores were not significantly different based on the order the participants experienced those conditions, indicating that there was not a significant learning effect.

A two-factor repeated measures ANOVA revealed that the number of call signs had a significant effect on RT ($F(1, 17) = 11.786$, $p = 0.003$, partial eta = .409), but there was no significant effect of PI (although it approached significance at $p = .073$) or the interaction on RTs, as shown in the right panel of Figure 4. A repeated measures single factor ANOVA with a Greenhouse-Geisser correction revealed that the RTs across trials were significantly different, ($F(1.7, 28.5) = 8.520$, $p = 0.002$, partial eta = 0.334). Post hoc tests using the Bonferroni correction determined that RTs in 5-PI (*M* = 3.338 *SD* = .342) were statistically significantly lower than RTs in 7-NP (*M* = 3.587, *SD* = .405) and 7-PI (*M* = 3.579, *SD* = .430). The RT for 5-NP (*M* = 3.425, *SD* = .316) was not significantly different from the others.

Additionally, an independent samples t-test indicated that RTs were significantly different based on the order participants experienced the 5 versus 7 call sign condition ($t(76) = 3.034$, $p = .003$) where those experiencing the 5-CS conditions first had a significantly higher RT (*M* = 3.601, *SD* = .376) than those who experienced the 7-CS conditions first (*M* = 3.352, *SD* = .349).

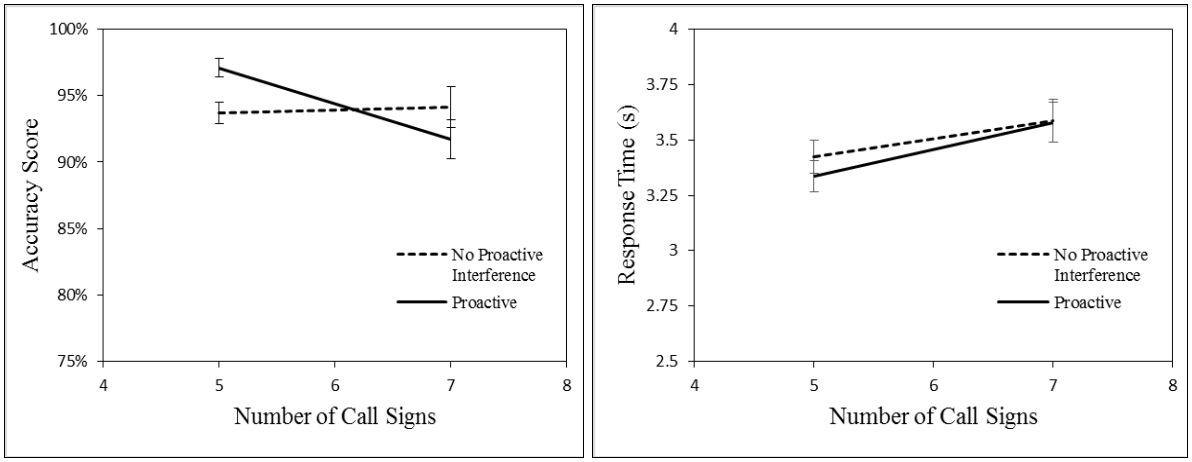

Figure 4. Interaction of number of call signs on accuracy scores for both PI conditions,(left panel) and the interaction of number of call signs on response times for both PI conditions (right panel).

A repeated measures ANOVA determined that there was no significant difference between workload or difficulty measures across all trials. Additionally, when asked "based on your experience today, how many call signs do you think you could monitor comfortably before you would begin missing time critical information?" after all experimental trials, participants responded with a mean of 5.86 (*SD* = 1.35). Responses ranged from 3 to 8 call signs.

Discussion

Overall, the results show that the participants' accuracy and response time was degraded as the number of call signs increased from 5 to 7, as expected. However, the results with respect to proactive interference differed from expected as accuracy and response time were not consistently degraded in the presence of proactive interference. Specifically, with respect to the accuracy scores, the 5 call sign PI condition was the highest scoring even though it was not the lowest taskload condition. A few possible explanations could be offered.

First, the workload-performance curve (similar to the Yerkes-Dodson Law) shows that high and low levels of workload result in low performance, but medium levels of workload result in higher performance (Teigen, 1994) creating an inverted-U shaped relationship. One potential explanation is that the workload was so low that the participants' performance did not reach its optimal level. This, however, was not supported by the reported workload and difficulty scores which did not significantly differ across the experimental conditions.

As it is necessary for the participants to be exposed to a set of call signs before these same call signs can induce proactive interference, another possible explanation stems from the need to present the PI conditions after the NP conditions. The results indicated that RT was influenced by whether the participants experience the 5 or the 7 call sign condition first, potentially indicating that the participants who experienced the 7 call sign condition first underwent a higher rate of learning than the participants who experienced the 5 call sign condition first. It is possible that negative effects of proactive interference were offset by learning effects within the current experiment.

Sampling error could have also contributed to the unexpected outcomes. For most variables, there was data from only 21 participants (due to missing data). Because of this small sample size, irregular data points could have been magnified in the results. Although the trials were kept to a short length, fatigue could have been a factor in this study, as some participants reported feelings of boredom. Additionally, there were a limited number of call signs used in this experiment, with only 13 call signs available for use in the trials. As a result, on trials where participants were supposed to remember 7 call signs, some reported that instead of listening for the call signs on the list, they listened for the ones not on the list since they believed (correctly) that there were fewer of those. Ideally, a new set of call signs would be used on each trial to prevent habituation.

Conclusion

The results of this study provide conflicting evidence about whether higher taskload conditions actually produce lower levels of performance. This study indicated that increasing the number of call signs from 5 to 7 reduced the participants' accuracy and increased their response time. However, the results do not support the hypothesis that performance will be reduced by proactive interference, a result which has multiple potential explanations including learning, workload, and sample bias effects. Further research is recommended which include additional task load levels (more call signs/PI conditions), more participants, less overlap in call signs between conditions, and potentially enhanced training. Data from this research could give insight into a relationship that exists among these variables.

Acknowledgements: The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the Department of the Air Force, Department of Defense, nor the U.S. Government. The authors would like to thank Colonel Anthony Tvaryanas for motivating the current study.

References

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science, 15(2),* 106-111.

Baddeley, A. (2000). The episodic buffer: a new component of working memory?. *Trends in Cognitive Sciences, 4(11)*, 417-423.

Baddeley, A.D. and Hitch, G. (1974) *Working memory*. In The Psychology of Learning and Motivation (Bower, G.A., ed.), pp. 48-79, Academic Press.

Blair, E. A., Rahill, K. M., Finomore, V., Satterfield, K., Shaw, T., & Funke, G. (2014, September). Best of Both Worlds Evaluation of Multi-Modal Communication Management Suite. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 58, No. 1, pp. 410-414). SAGE Publications.

Engle, R. W., & Oransky, N. (1999). Multi-store versus dynamic models of temporary storage in memory. *The Nature of Cognition*, 515-555.

Finomore, V.S., Popik, D.K., Castle, C.E., & Dallman, R.C. (2010). Effects of network-centric multi-modal communication on a communication-monitoring task. *Proceedings of the Human Factors and Ergonomics Society, 54,* 2125- 2129.

Finomore, V., Stewart, J., Singh, R., Raj, B., & Dallman, R. (2012). Demonstration of advanced multi-modal, network-centric communication management suite. *Proceedings of the Interspeech, 13.*

Finomore, V., Satterfield, K., Sitz, A., Castle, C., Funke, G., Shaw, T., & Funke, M. (2012). Effects of the multi-modal communication tool on communication and change detection for command and control operators. *Proceedings of the Human Factors and Ergonomics Society, 56, 1464-1465*.

Ishihara, S. (1980). *Ishihara's design charts for colour-blindness of unlettered persons*. Kanehara & Company.

Kane, M. J., Bleckley, M. K., Conway, A. R., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General, 130(2)*, 169.

Kane, M. J., & Engle, R. W. (2000). Working-memory capacity, proactive interference, and divided attention: limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26(2)*, 336.

Kumar, S., Joseph, S., Pearson, B., Teki, S., Fox, Z. V., Griffiths, T. D., & Husain, M. (2013). Resource allocation and prioritization in auditory working memory. *Cognitive Neuroscience, 4(1)*, 12-20.

Logarithmic Near Visual Acuity Chart 2000 "New ETDRS" Charts (2011). Precision Vision.

May, C. P., Hasher, L., & Kane, M. J. (1999). The role of interference in memory span. *Memory & Cognition, 27(5)*, 759-767.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review, 63(2),* 81.

Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience, 139(1),* 23-38.

Santana, L., Langhals, B., Miller, M., & Finomore, V. (2013). Does supplementary computer generated cueing enhance controller efficiency in a congested communication environment? *Proceedings of the International Symposium on Aviation Psychology, 17,* 226-231.

Sweller, J (1988). "Cognitive load during problem solving: Effects on learning". *Cognitive Science 12(2)*: 257–285.

Tattersall, A. J., & Foord, P. S. (1996). An experimental evaluation of instantaneous self-assessment as a measure of workload. *Ergonomics*, *39*(5), 740-748.

Teigen, K. H. (1994). Yerkes-Dodson: A law for all seasons. Theory & Psychology, 4(4), 525-547.

III. COGNITIVE ABILITY PREDICTORS OF MULTIAIRCRAFT CONTROL (MAC) TASK PERFORMANCE

Kelly M. Amaddio, Michael E. Miller, John Elshaw
Air Force Institute of Technology
Dayton, Ohio

ABSTRACT: Several laboratory studies have been conducted on individual differences in working memory capacity and attention; however few studies have applied these principles to a real-world example. This study seeks to explore the relationship between working memory, attention, and performance on a multiaircraft control radio communication task by using data from the Operations Word Span test, Attention Control Scale, and GRE scores. Hierarchical linear modeling was used to determine the relationships among the variables.

INTRODUCTION

Remotely Piloted Aircraft (RPAs) are becoming increasingly important across many industries. In the military context, this technology has significantly enhanced the gathering of intelligence, surveillance and reconnaissance (ISR) without risking bodily injury to military and border patrol personnel (New York Times, 2013). Further, this technology is expected to be useful in many civilian contexts, including meteorology, wildlife preservation, agriculture, search and rescue, and logistics (Handwerk, 2013). RPAs differ from manned aircraft along many dimensions, but a difference important in the current research stems from the fact that since an operator is not required in each aircraft, it is possible that one operator may control many aircraft simultaneously. For example, transit operators may be employed to simultaneously pilot multiple semi-autonomous aircraft between an airbase and an area of operations where aircraft control would be subsumed by individual pilots to avoid operator overload (Colombi et al., 2012).

In such a scenario, the transit pilot may have to monitor and respond to radio communications for each of several aircraft, potentially resulting in a large throughput of radio calls even though the number of radio calls for an individual aircraft is sparse. Additionally, there is a concern that proactive interference may occur when the transit pilots transfer aircraft to other pilots, but still hear the previous aircraft specific radio calls. Several principles related to working memory, interference, and attention are important to the analysis of this issue.

The ability of an operator to listen to and respond appropriately to radio traffic which contains references to the call signs of the aircraft they are controlling, as well as other entities, is likely to be constrained by their available working memory. Working memory is involved in storing and manipulating information for short term use in tasks like reasoning and comprehension (Baddeley & Hitch, 1974). A common model of working memory that has been proposed by Baddeley (2000) contains a set of subsystems, including the central executive which controls attention between the visuospatial sketchpad, episodic buffer, and phonological loop subsystems. The visuospatial sketchpad manipulates visual images while the phonological loop is responsible for storing and replaying words and sounds. The episodic buffer temporarily stores and integrates multimodal information and relays information between the visuospatial sketchpad and phonological loop. The auditory component of this model is important to the current study because participants are asked to listen and respond to a series of aircraft radio calls.

Although significant research has been conducted on visual working memory, auditory working memory has garnered less attention. Perhaps this emphasis is due to the prevalence of traditional auditory working memory measures that assess performance in a binary fashion where the item is either remembered or not, while sounds contain many perceptually distinct attributes, including pitch, language, semantic meaning, rhythm, etc. that need to be considered. Considering this, Kumar et al. (2013) attempted to test auditory working memory over a continuous scale by using sequences of tones in different lengths. Participants listened to a sequence of tones and afterwards had to adjust a dial to replicate a specific tone that they heard. They found that increasing the number of tones held in working memory reduced the precision of the memories, much like what is found in visual working memory (Alvarez & Cavanagh, 2004). However, the task applied in this research differs significantly from the application of current interest as the pilot will assign meaning to each call sign, rather than remember the specific tone. Therefore, it is likely that in the current task, each call sign will increase the operator's cognitive load, increasing the mental effort used to maintain information in working memory (Sweller, 1988). Therefore, limits in individuals' information processing capacity will limit the number of call signs to which the pilot can retain and respond. It is therefore, reasonable that an operator may be able to retain and recall approximately 7±2 items (Miller, 1959).

One of the primary functions of working memory is to navigate the effects of proactive interference (Kane & Engle, 2000) where timely information replaces less recent information to reduce the likelihood of confusion. Therefore, effective working memory will suppress memory of outdated information to prevent it from interfering with the encoding of new information. Proactive interference has been shown to affect performance on working memory tasks. May, Hasher, and Kane (1999) found that performance on a working memory span test was improved when measures were taken to prevent proactive interference (e.g., temporally separating trials).

Kane and Engle (2000) found that individuals with low working memory spans showed greater susceptibility to proactive interference under low cognitive load conditions, but under high cognitive load conditions, both high and low working memory span individuals showed equal levels of proactive interference. These results suggest that attention plays a role when encoding and retrieving memories as working memory is usually measured by span tasks that require the individual to simultaneously process and remember verbal information.

A competing model to working memory has been suggested by Engle and Oransky (1999) who propose that controlled attention is the mechanism by which working memory functions. They describe controlled attention as the "ability to effectively maintain stimulus, goal, or context information in an active, easily accessible state in the face of interference, to effectively inhibit goal-irrelevant stimuli or responses, or both" (Kane, Bleckley, Conway, & Engle, 2001, p.18). Neurological evidence shows that different information (sensory, semantic, etc.) is stored in different areas of the brain (Postle, 2006) suggesting that working memory should be seen as directing attention towards different memory codes stored in long term memory. Although these models of working memory have different implications for the design of interfaces to support the simultaneous control of multiple aircraft by a single individual, they all support the view that the operator's attention must be divided between the visuospatial tasks necessary to control the aircraft, processing of audio call signs and the integration of this information, as described in multiple resource theory (Wickens, 1981).

This study seeks to understand the ability of an operator to respond to numerous call signs from among a number of distracters to understand performance in this multi-aircraft communications task. This relationship will be explored, both in the absence and presence of proactive interference. Additionally, the research explores the relationship between several standard measures of cognitive performance and the performance of individuals in the multi-aircraft control radio communications task. As the current literature does not appear to connect basic laboratory working memory tests with real-life applications of verbal recall, it is hoped that the current study will not only provide information regarding operator performance while monitoring multiple call signs in a multi-aircraft communications task but provide information regarding the relationship between standard working memory tasks and performance of the multi-aircraft communications task.

Based on previous findings, it is predicted that there will be a positive relationship between cognitive ability and performance in the current study. The measures that will be used to represent cognitive ability are the GRE (Graduate Record Examination), Attention Control Scale, and Operations Word Span test. Additionally, the relationship between cognitive load and performance will be moderated by proactive interference.

Method

Participants

Twenty one (5 female and 16 male) volunteers with ages between 22 and 44 (M = 27.75, SD = 4.96) participated in the study. Participants were required to have a visual acuity of 20/30 or better as determined using the Logarithmic Near Visual Acuity Chart 2000 ("New ETDRS" Charts, 2011)) and normal color vision, using pseudo-isochromatic plates (Ishihara, 1980). There was no educational requirement, however all participants had or were seeking a graduate degree and most participants were engineering students. All participants read and agreed to an informed consent document. They were then assigned a participant number and no personally identifiable information was retained per Institutional Review Board Protocol.

Apparatus

The experiment was conducted in a 6ft x 6ft cubicle in a quiet laboratory to minimize distractions. The experimental setup consisted of Bose AE2w headphones and a laptop to present the call signs using the Multi-modal Communication (MMC) software (Finomore, Popik, Castle, & Dallman, 2010). Participants were also given a wireless number keypad, a clipboard containing a number grid, and a clipboard containing the list of call signs for each trial. The list of call signs was attached to the left wall to the cubicle slightly above eye level. The placement of this list required the participant to actively turn their head to view the list of call signs.

The Multi-modal communication program (MMC) is an Air Force Research Laboratory developed multi-modal, network-centric communication management suite developed to aid Command and Control operators in increasing communication intelligibility and reduce mental workload. This tool combines several features designed to improve the performance of the users, including spatial audio, speech transcription, data capturing and playback, chat messages, and automatic keyword highlighting (for full description of the MMC tool see Finomore et al., 2012). Additionally, this tool has been used extensively as a research tool to evaluate a variety of communication effectiveness questions (Blair, Rahill, Finomore, Satterfield, Shaw, & Funke, 2014; Finomore, et al. 2010; Finomore, Stewart, Singh, Raj, & Dallman, 2012; Finomore, Satterfield, Sitz, Castle, Funke, Shaw, & Funke, 2012; Santana, Langhals, Miller & Finomore, 2013). This experiment utilized monaural sound, a chat window to prompt the participant to enter the appropriate code, and the logging function to record the participants' inputs.

Cognitive Performance Measures

Three separate measures of cognitive ability were selected as potential predictors of individual performance. These measures included the Operations Word Span Test, the Attention Control Scale and the results of the Graduate Record Examination.

Operations Word Span Test

The Operations Word Span test was adapted from Turner and Engle (1989). The original test contained a series of mathematical expressions followed by a to-be-remembered word. The number of mathematical operations and words in a trial increased from two to five in the original experiment with each number of mathematical expression-word pairs being administered three times. In this test, the mathematical operations were performed on integer numbers less than ten. Each expression contained an initial multiplication or division operation in parenthesis followed by an addition or subtraction of a single number and stated a potential answer e.g. [(3 x 2) + 1 = 7]. The potential answer was correct approximately half of the time. Successful participants were required to state that the answer was "correct" or "incorrect" with an accuracy of 80% or greater. Following the mathematical expression, a single syllable, to-be-remembered word was presented. In the original test, the words were one-syllable concrete nouns selected from a published frequency norm list. The participant said the word aloud when it was displayed and the next mathematical expression – word pair was presented. After the appropriate number of mathematical expression – word pairs were displayed, the screen was blanked and the participant was asked to repeat each of the to-be recalled words.

Within the current experiment, this task was altered slightly. Specifically, participants were asked to respond to between 2 and 7 mathematical expression-word pairs, with each trial administered twice using a PowerPoint slide deck. Participants in this study pressed "Enter" on the keyboard to advance the slides on their own pace. Response data was collected by the investigator using a paper score sheet. The words in this experiment were two-syllable concrete nouns that were not within the visual field of the subject during testing; for example, the word "keyboard" was not used to the presence of a keyboard on the desk. Additionally, special care was taken to ensure that none of the words in a list belonged to the same category e.g. food, animals, sports, etc. Two-syllable words were used because the call signs in the MMC experiment also contained two syllables. Before the experimental trials started, participants were given a two-word warm-up trial and permitted to ask any questions. It should also be noted that this task was administered to participants during the midpoint break in the experiment.

In the analysis, the results of the OSPAN test were scored by adding up all of the correct word responses from each trial (regardless of trial number) into an aggregate score out of 54 possible points. It is predicted that this score is positively related to performance on the current task.

Attention Control Scale

The participants completed the Attention Control Scale (ACS) questionnaire (Derryberry & Reed, 2002) after completing the four experimental trials. This questionnaire consisted of twenty 4-point Likert-type questions that contained two subscales that measured an individual's ability to focus and shift attention. The psychometric properties of this scale were investigated and the results indicated a moderate test-retest reliability and high internal consistency reliability (Cronbach's alpha = .88) (Fajkowska & Derryberry, 2010). Responses on the current study achieved a Cronbach's alpha of .77. It is predicted that both attention focusing and shifting abilities are positively related to performance on the multiaircraft control task.

Graduate Record Examination (GRE)

Working memory has been shown to be strongly correlated with performance on several measures of cognitive ability, including the verbal Scholastic Aptitude Test (SAT) (Daneman & Carpenter, 1980) and reading comprehension. The Graduate Record Examination (GRE) is a computer based test used to provide a standardized score for graduate school applicants. It contains three sections: verbal reasoning, quantitative reasoning, and analytical writing (About the GRE ® Revised General Test). This test is used as a proxy for cognitive ability and has been shown to be a valid predictor of graduate school grade point average, examination scores, publication counts, and faculty ratings (Kuncel, Hezlett & Ones, 2001). Since the Graduate Record Examination (GRE) is designed to measure verbal and quantitative reasoning, it would be reasonable to assume that may also be predictive of scores on a variety of other cognitive measures. Furthermore, it is predicted that GRE scores are positively related to performance on the multiaircraft control task. It should be noted that this measure was not collected specifically for this experiment but was acquired from the participants' official student records and analytical writing scores were not analyzed.

Experimental Procedure

Participants were randomly assigned to one of two groups based on their subject number. They were given a quick explanation of the software and task, and then completed a three minute practice session in which they were responsible for three call signs. This practice session permitted participants to experience the software and practice the task to reduce learning effects during the experiment. Although a hearing test was not administered, participants were encouraged to adjust the volume to their comfort level during the practice session.

Based on their group, participants were asked to attend to either 5 or 7 call signs (out of 13 possible call signs) during each of four 8-minute experimental trials. The trials were counterbalanced to offset any potential learning effect. Participants in Group 1 were assigned five call signs for the first two trials and seven for the second two trials. The participants in Group 2 were assigned seven call signs for the first two trials and five call signs for the second two trials. Each 8-minute trial contained 100 radio calls that were evenly spaced 5 seconds apart. Approximately 50 radio calls were critical and an equal number were distracters. This ratio was not revealed to the participants. During the second and fourth trials, 20 of the distracters were selected to introduce proactive interference (PI) as they were selected from among the critical call signs in the previous trial. The order of the radio calls and calls signs was randomized. Table 2 presents the trials and the critical and PI call signs for Group 1. The call signs were the same for Group 2, except they experienced the 7 call sign conditions first. The scenarios will be referred to as 5-NP (5 call signs, no PI condition), 5-PI (5 call signs, PI condition), 7-NP (7 call signs, no PI condition), and 7-PI (7 call signs, PI condition). Table 2 also shows that Trials 2 and 3 were separated by a break that required approximately 15 minutes during which the participants completed the Operations Word Span Test followed by a period during which they were encouraged to leave the lab to obtain a drink of water and visit the restroom.

Table 2.
Call signs experienced by Group 1 during each trial. Call signs which were employed to induce PI during Trials 2 and 4 are shown in Bold-Italics for Trials 1 and 3.

Participant Group	**Trial 1** (5-NP)	**Trial 2** (5-P)	Break (15 minutes)	**Trial 3** (7-NP)	**Trial 4** (7-P)
1	Laker Hopper Arrow ***Charlie*** ***Gringo***	Laker Hopper Arrow Tiger Eagle	Operations Word Span test followed by a break	Charlie Gringo Laker Raptor Viking ***Arrow*** ***Tiger***	Charlie Gringo Laker Raptor Viking Thunder Cobra

The participants were instructed to listen for the commands that contained their call sign. Each radio call began with the word “Ready”, which was proceeded by a call sign and a command containing a grid coordinate, for example, “Ready Charlie go to blue one now.” The color indicates a column in the grid and the number represents a row in the grid. The grid location would then contain a number. For critical call signs, the participants found the space on the grid that corresponded with the command, and typed the number from the grid location into the MMC chat window. For example, when the participant heard “Ready Charlie go to blue one now,” if the participant was responsible for “Charlie” during that trial (Charlie would be on their list of call signs), they would be expected to find the “blue 1” spot on the grid and type that number from that grid location on the keypad. If the participant heard a call sign that was not on their list, they were instructed to type a zero into the chat window. Also, if for some reason they were not sure whether they were responsible for a specific call sign, they were instructed to type a zero as the participants were told that it was more important to answer only the call signs they were responsible for rather than responding to another operator’s call sign. The randomized numbers on the grid were between 10 and 99. Participants were given as much time as they needed to memorize the call sign list before every trial and were instructed to only look at the list of call signs if they forgot them during the trial. The number of times they referred to the call sign list (Looks at board) was recorded by the investigator for every trial.

To avoid habituation to call signs and voices, certain measures were taken. First, the list of critical call signs on the clipboard was shuffled for each trial, making it difficult to memorize the call signs from the call sign list. All trials contained a different random radio call order, different call signs, and called for different grid locations. Additionally, a new number grid was used for each trial. Finally, each of the 13 call signs were recorded with different voices, permitting different voices to make radio calls for every call sign so that the participant could not ignore or attend to a certain call sign based on the speaker. During the experiment, each participant heard 12 difference individual's voices and 13 different call signs.

Performance Measures

Data was collected during all trials using the logging function in the MMC software to compare user performance under different conditions and different times. Numerical responses submitted by the participants on the MMC task were evaluated for accuracy and RT. For each trial, the accuracy score was calculated by dividing the number of correct responses by the total number of radio calls and multiplying by 100%. Additionally, a PI accuracy percentage correct score was determined. This value is the number of correct responses given for the PI call signs divided by the total number of radio calls expected to introduce PI for the 5-PI and 7-PI conditions. Finally, the average of the participant's RTs were calculated for each trial as the average time lapse between the end of the radio call and the "Enter" key press after the numerical numeric response was typed.

After each trial, participants were asked to respond to two 5-point Likert Scale questions. The first question was an instantaneous self-assessment of workload (Tattersall & Foord, 1996). The second question regarded the perceived difficulty of the trial and included the descriptions (1= very easy, 2 = easy, 3 = neutral, 4 = difficult, and 5 = very difficult). After the last trial, participants were asked to self report the number of call signs they felt like they could reliably monitor (CS Monitor). Finally, they completed the questionnaire associated with the attention control scale.

Data Analysis

Considering the repeated measures nature of this research, hierarchical linear modeling (Hoffman, 1997) was applied to analyze the data. This type of analysis accounts for the fact that the data contains non-independent, hierarchically structured samples and therefore violates the assumptions of traditional regression. Additionally, linear regression neglects the shared variance among the groups of data in this analysis (Woltman, Feldstain, MacKay & Rocchi, 2012). This analysis included five level 2 variables, including GRE scores, average RT, CS monitor, age, and gender. Additionally, seven level 1 variables, including perceived workload and difficulty, number of call signs, amount of times participants referenced the call sign list, accuracy scores, and RTs) were included in this analysis.

Results and Discussion

A two-factor repeated measures ANOVA revealed that there was a significant main effect of the number of call signs as well as the interaction between the number of call signs and the presence of PI on accuracy scores on the MMC task ($F(1, 19) = 7.631$, $p = 0.012$). Proactive interference moderates the relationship between the number of call signs and accuracy scores such that the relationship is weaker for individuals when no proactive interference is present, as shown in Figure 5. In other words, the relationship between number of call signs and accuracy score is stronger when proactive interference was present.

The interaction was further analyzed by applying a single factor repeated measures ANOVA. This analysis revealed that the accuracy scores were significantly different across trials, $F(2.280, 43.31) = 4.307$, $p = 0.016$, partial eta squared = 0.19. Post hoc tests using the Bonferroni correction determined that scores in the 5-PI condition (M = 97.11%, SD = 3.75%)

were statistically higher than scores in the 5-NP condition (*M* = 93.70%, *SD* = 3.16%) and 7-PI conditions (*M* = 91.73%, *SD* = 6.48%). The scores for 5-NP, 7-NP (*M* = 94.14%, *SD* = 6.44%), and 7-PI were not significantly different from one another. Therefore, we can conclude that the highest scoring condition occurred when the participants were tasked with 5 call signs in the PI condition. A paired samples t-test indicated that PI accuracy scores were not significantly different between 5-PI (*M* = 95.29%, *SD* = 12.63%) and 7-PI (*M* = 90.25%, *SD* = 15.27%). Additionally, an independent samples t-test showed that accuracy scores were not significantly different based on the order the participants experienced those conditions, indicating that there was not a learning effect. Since one of the goals of the current study was to evaluate the effects of proactive interference on performance, a more robust counterbalancing technique (like Latin Squares) was not used because the proactive interference condition always had to occur in trials 2 and 4. Thus, it was impossible to prevent learning from occurring from trials 1 to 2 and 3 to 4.

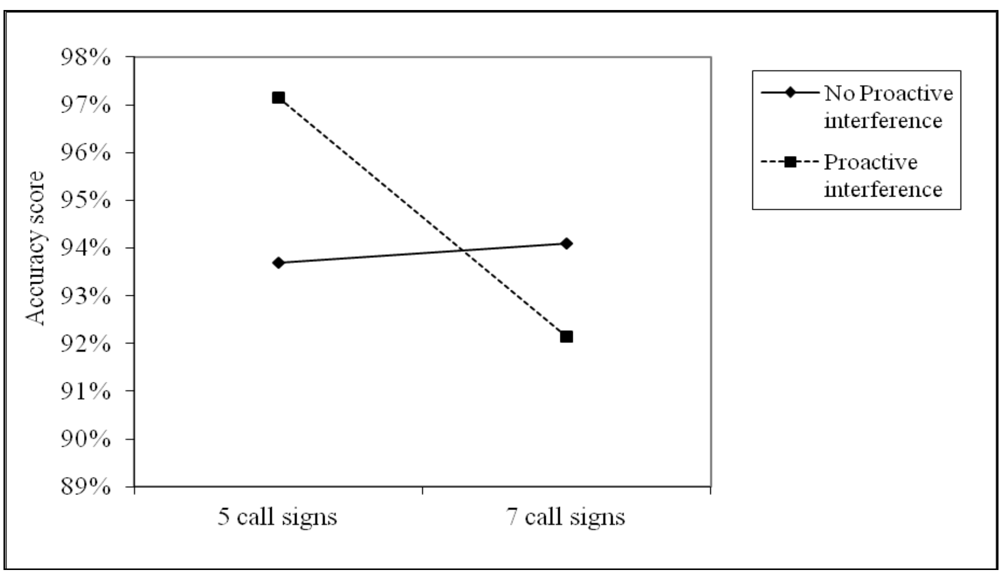

Figure 5. Proactive interference moderation between the number of call signs and accuracy scores. Note: a simple slopes test was conducted on both high and low levels of the moderator and the slope of both lines was significant ($p < .01$).

Hierarchical linear modeling was used to determine the variables which predicted accuracy scores on the multiaircraft control task. The results of this analysis are shown in Table 3. RT, and the number of call signs have significant negative Beta values, indicating that as the reaction time and the number of call signs increase, accuracy decreases. Quantitative GRE score, "Looks at board", and the number of call signs were also shown to be significant predictors of accuracy. This model shows that, RT is the strongest predictor of accuracy scores and quantitative GRE, attention focusing (nearly significant) and looks at board are similar levels of predictors. Verbal GRE, OSPAN, and attention shifting were weak and nonsignificant predictors of accuracy.

Table 3
Summary of Hierarchical Linear Regression Analysis for Variables Predicting Accuracy Scores (N = 21)

Variable	γ	*p*
RT	-0.53	< .001
Quantitative GRE*	.25	0.035
Attention Focusing	-0.256	.078
Looks at board*	0.24	.039
Number of CS	-0.17	.04
Verbal GRE*	.059	.632
Attention Shifting	-0.003	.979
OSPAN	.017	.893

Note. Values containing an asterisk were obtained by controlling for response time. Standardized scores were used on all variables except Number of CS.

A similar analysis was conducted to understand the variables which predict RT. The hierarchical regression indicated that the number of CS ($\gamma = .179$, $p < .001$) and looks at board ($\gamma = .009$, $p = .023$) had a small significant effect on RT, but no other cognitive tests were significantly predictive of RT on the MMC task. The response time slowing as a result of the addition of more CS is an expected outcome because the added task load causes the individual to process more information (Barrouiller, et al., 2007).

Several correlations were found between the variables in this study. There were both expected and unexpected correlations. As the number of call signs increased, RT increased and accuracy decreased, which is consistent with the speed/accuracy trade-off seen in many cognitive tasks (Wickelgren, 1977). OSPAN was positively correlated with GRE quantitative and attention shifting, which is an expected outcome, since attention control is considered to be a mechanism by which working memory effectively operates (Engle and Oransky, 1999) and the OSPAN task requires individuals to shift attention from one set of words to another within seconds. Since the GRE is a measurement of cognitive ability and working memory capacity is related to other measures of cognitive ability, like the SAT (Daneman & Carpenter, 1980). It is interesting, however, that OSPAN is not significantly correlated with both verbal and quantitative GRE scores. Since the participants were asked to remember sets of words during the OSPAN test, it is surprising that it is not positively correlated with verbal GRE scores as well. Additionally, attention focusing and attention shifting were positively correlated with a medium effect size. This is expected because they measure different but complimentary attention mechanisms. Another unexpected correlation that emerged was the negative relationship between attention focusing and both GRE measures (verbal and quantitative) and accuracy. Further, while not statistically significant, a negative correlation also exists between Attention Shifting and OSPAN performance. These correlations have a small-medium effect size. The positive relationship between OSPAN and looks at board is unexpected because those with high working memory span should reference the call sign list less. There could be a mediating personality variable causing this effect.

Table 4
Correlations Between Experimental Variables

Variable	1	2	3	4	5	6	7	8
1) RT	-							
2) Accuracy	-.420***	-						
3) Number of CS	.246*	-.232*	-					
4) Attn Focus	.006	-.223*	-.014	-				
5) Attn Shift	-.083	-.007	.004	.327**	-			
6) GRE Verbal	.004	.106	.011	-.230*	-.120	-		
7) GRE Quant	.007	.211	.003	-.317*	.155	.623**	-	
8) OSPAN	-.055	.040	.019	-.225*	.223*	-.092	.305**	-
9) Looks at Board	-.053	.268*	.085	.047	-.079	-.092	.134	.257*

Note: Correlations are given *, **, or ***, depending on whether the p value is less than .05, .01, or .001.

Unfortunately the results from the OSPAN test and attention shifting were not highly predictive of accuracy scores on the MMC radio communications task (although attention focusing approached significance in negatively predicting accuracy scores). The discrepancy between the findings from past studies and the current results could be caused by a number of factors. First, because each trial of the MMC task was eight minutes long, this is a very different measure than the OSPAN test where participants are directed to recall a list of words within seconds of learning them. Additionally, the current experiment allowed participants to reference the list of words when they thought they may have forgotten them. The presence of interference is a factor that strongly affects performance on the MMC task, as shown by the moderation.

A significant limitation on the analysis of this study is the small sample size. According to Cohen (1992), to find a significant result in a correlation analysis at the $\alpha = .05$ level, there would need to be 783, 85 or 28 participants for a small, medium, or large effect, respectively. For a multiple regression, there would need to be at least 481, 67, or 30 participants for a small, medium, or large effect, respectively. This insufficiency in sample size may cause a type II error to occur because the power isn't large enough to detect a significant result. Additionally, when analyzing regression statistics with a standardized beta value, since the beta value is determined by the standard deviation, any extreme data points could drastically change the results. The small sample size could also be causing the unexpected correlations, although the negative correlation between attention focusing and GRE scores, OSPAN, and accuracy deserves extra analysis.

Additionally sampling bias may have occurred since this was a sample of convenience. All participants were graduate students or higher, all were in the Air Force or worked for the Department of Defense, most were engineers, and the majority of the engineers were civil engineers. This resulted in a homogeneous population whose scores may not be representative of the population. Additionally, the individuals in the sample could have similar personality traits that confounded certain results. This could also explain why the current results do not support the findings in earlier research. Compared to an earlier sample obtained by Fajkowska & Derryberry (2010) who found a mean score of 54.49 with a standard deviation of 15.02 on the Attention Control Scale in a sample of 218 undergraduate students. The current sample had a similar mean of 53.19 but a standard deviation of 6.43, meaning there is less variance in the current sample. Additionally, the GRE is has an overall mean score of 150.54 in the verbal and 152.14 quantitative categories (ETS, 2014), but the mean scores in this study are 153.35 and 155.20, respectively.

Conclusion

The current study found that PI moderates the relationship between cognitive load and performance on the multiaircraft control task where PI has a stronger effect at higher cognitive loads. Because of this, measures should be taken to reduce PI in certain tasks. It was also found that certain measures of cognitive ability like quantitative GRE scores and attention focusing are predictive of performance on a multiaircraft radio communications task, but more research needs to be conducted to find stronger predictors. Additionally, different measures of cognitive ability were found to be mildly related, but a larger and more diverse sample would be useful to show a stronger relationship.

References

About the GRE® revised General Test. (n.d.). Retrieved February 5, 2015, from http://www.ets.org/gre/revised_general/about?WT.ac=grehome_greabout_b_130807.

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science, 15(2),* 106-111.

Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(3), 570.

Baddeley, A. (2000). The episodic buffer: a new component of working memory?. *Trends in Cognitive Sciences, 4(11)*, 417-423.

Baddeley, A.D. and Hitch, G. (1974) *Working memory*. In The Psychology of Learning and Motivation (Bower, G.A., ed.), pp. 48-79, Academic Press.

Blair, E. A., Rahill, K. M., Finomore, V., Satterfield, K., Shaw, T., & Funke, G. (2014, September). Best of Both Worlds Evaluation of Multi-Modal Communication Management Suite. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 58, No. 1, pp. 410-414). SAGE Publications.

Cohen, J. (1992). A power primer. *Psychological bulletin*, *112*(1), 155.

Colombi, J., Miller, M., Schneider, M., McGrogan, J., Long, D., & Plaga, J. (2012). Predictive mental workload modeling for semi-autonomous system design: Implications for systems of systems. *Systems Engineering, 15*(4), 448-460.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450-466.

Derryberry, D., & Reed, M. A. (2002). Anxiety-related attentional biases and their regulation by attentional control. *Journal of abnormal psychology*, *111*(2), 225.

Engle, R. W., & Oransky, N. (1999). Multi-store versus dynamic models of temporary storage in memory. *The Nature of Cognition*, 515-555.

ETS (2014). *Performance statistics on the GRE revised general test.* Retrieved February 10, 2015 from https://www.ets.org/s/gre/pdf/gre_guide_table1a.pdf.

Fajkowska, M., & Derryberry, D. (2010). Psychometric properties of Attentional Control Scale: The preliminary study on a Polish sample. *Polish Psychological Bulletin*, *41*(1), 1-7.

Finomore, V.S., Popik, D.K., Castle, C.E., & Dallman, R.C. (2010). Effects of network-centric multi-modal communication on a communication-monitoring task. *Proceedings of the Human Factors and Ergonomics Society, 54,* 2125- 2129.

Finomore, V., Stewart, J., Singh, R., Raj, B., & Dallman, R. (2012). Demonstration of advanced multi-modal, network-centric communication management suite. *Proceedings of the Interspeech, 13.*

Finomore, V., Satterfield, K., Sitz, A., Castle, C., Funke, G., Shaw, T., & Funke, M. (2012). Effects of the multi-modal communication tool on communication and change detection for command and control operators. *Proceedings of the Human Factors and Ergonomics Society, 56, 1464-1465*.

Handwerk, B. (2013, December 2). 5 Surprising Drone Uses (Besides Amazon Delivery). Retrieved February 12, 2015, from http://news.nationalgeographic.com/news/2013/12/131202-drone-uav-uas-amazon-octocopter-bezos-science-aircraft-unmanned-robot/.

Hofmann, D. A. (1997). An overview of the logic and rationale of hierarchical linear models. *Journal of management*, *23*(6), 723-744.

Ishihara, S. (1980). *Ishihara's design charts for colour-blindness of unlettered persons*. Kanehara & Company.

Kane, M. J., Bleckley, M. K., Conway, A. R., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General, 130(2)*, 169.

Kane, M. J., & Engle, R. W. (2000). Working-memory capacity, proactive interference, and divided attention: limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26(2)*, 336.

Kumar, S., Joseph, S., Pearson, B., Teki, S., Fox, Z. V., Griffiths, T. D., & Husain, M. (2013). Resource allocation and prioritization in auditory working memory. *Cognitive Neuroscience, 4(1)*, 12-20.

Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2001). A comprehensive meta-analysis of the predictive validity of the graduate record examinations: implications for graduate student selection and performance. *Psychological bulletin*, *127*(1), 162.

Logarithmic Near Visual Acuity Chart 2000 "New ETDRS" Charts (2011). Precision Vision.

May, C. P., Hasher, L., & Kane, M. J. (1999). The role of interference in memory span. *Memory & Cognition, 27(5)*, 759-767.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review, 63(2),* 81.

New York Times (2013), "Obama's Speech on Drone Policy," retrieved from http://www.nytimes.com/2013/05/24/us/politics/transcript-of-obamas-speech-on-drone-policy.html?pagewanted=all&_r=0.

Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience, 139(1),* 23-38.

Sweller, J (June 1988). "Cognitive load during problem solving: Effects on learning". *Cognitive Science 12 (2)*: 257–285.

Tattersall, A. J., & Foord, P. S. (1996). An experimental evaluation of instantaneous self-assessment as a measure of workload. *Ergonomics*, *39*(5), 740-748.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of memory and language*, *28*(2), 127-154.

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta psychologica*, *41*(1), 67-85.

Wickens, C. D. (1981). *Processing Resources in Attention, Dual Task Performance, and Workload Assessment* (No. EPL-81-3/ONR-81-3). Illinois Univ at Urbana Engineering-Psychology Research Lab.

Woltman, H., Feldstain, A., MacKay, J. C., & Rocchi, M. (2012). An introduction to hierarchical linear modeling. *Tutorials in Quantitative Methods for Psychology*, *8*(1), 52-69.

IV. Conclusions and Recommendations

Chapter Overview

This chapter is an overview of the current research, including the answers to the investigative questions stated in Chapter 1, the significance of the research, and recommendations for future research.

Research Overview

This research investigated the cognitive aspects of multiaircraft control. The experiment focused on the relationship between measures of cognitive ability and performance on a radio communications task. It also investigated the effect of cognitive demand and proactive interference on performance.

Answers to Investigative Questions

Question 1: **How is performance on a radio communications task related to cognitive demand (number of call signs and presence of proactive interference)?**

Proactive interference moderates the relationship between the number of call signs and accuracy scores such that the relationship is weaker for individuals when no proactive interference is present. In other words, performance measured in terms of accuracy and response time decreased as the number of all signs is increased from 5 to 7 when proactive interference was present.

Question 2: **Is an individual's level of attention control (attention focusing/attention shifting) related to their performance on a radio communications task in which the operator is monitoring multiple call signs?**

Each individuals' responses on the Attention Control Scale questionnaire were analyzed in relation to their accuracy scores and response time on the MMC task. Accuracy scores were significantly negatively correlated with attention focusing (with a small effect size). Attention focusing was found to almost significantly predict accuracy scores, once again with a negative relationship. This finding should be investigated further, as it is unexpected that a negative correlation would exist between these measures of attention and accuracy in this task. Attention shifting was not correlated with accuracy scores. Neither attention control subscale were related to response time. The influence of the particular participant sample involved in this study should also be considered when answering this question. The participants in this study belonged to a relatively homogeneous group where all were graduate students or higher and most were engineers and military personnel. This sample population may not be representative of the population and this could explain why some of the results were inconsistent with earlier findings and hypotheses. When compared with a different sample (Fajkowska & Derryberry, 2010) which had a mean Attention Control Scale score of 54.49 and standard deviation of 15.02, the current sample had a mean of 53.19 but a standard deviation of 6.43, meaning there is less variance in the current sample.

***Question 3:* Is an individual's working memory capacity related to their performance on a radio communications task involving multiple call signs?**

The results indicated that an individual's working memory capacity (tested by the Operations Word Span Test) was not found to be significantly correlated with accuracy and response time. Additionally, the hierarchical linear modeling results revealed that working memory capacity was not predictive of accuracy scores or response times. Although this finding is unexpected, it might be explained by the difference in this applied radio communication task where participants were able to refresh their memory as needed by referring to the call sign list throughout the experiment. This differs significantly from prior laboratory studies where participants were typically given prose recall tasks in which they recalled a list of words within seconds of memorizing them (Conway et al., 2005; Kumar et al, 2013).

Question 4: **What aspects of an individual are the most predictive of performance on a radio communications task: working memory capacity, attention control, GRE scores, etc.?**

The model shows that response time is the strongest predictor of accuracy scores and quantitative GRE, attention focusing (nearly significant), and looks at board are similar levels of predictors. Verbal GRE, OSPAN, and attention shifting were weak and non significant predictors of accuracy. None of the cognitive factors evaluated were determined to be predictive of response time. This could have occurred because an individual may have a high cognitive ability (high GRE scores, large working memory capacity) but may not be proficient at finding numbers and typing them quickly. Additionally, there may be other factors that better predict response time, like certain personality traits. An individual's personality could partially determine how they tackle the problem and whether they value accuracy or response time more.

Figure 6 shows the causal diagram that was suggested by the results of the current study. As indicated by the figure, not all of the predicted results were supported. The perceived workload and difficulty values were not related to the experimental conditions or the performance scores.

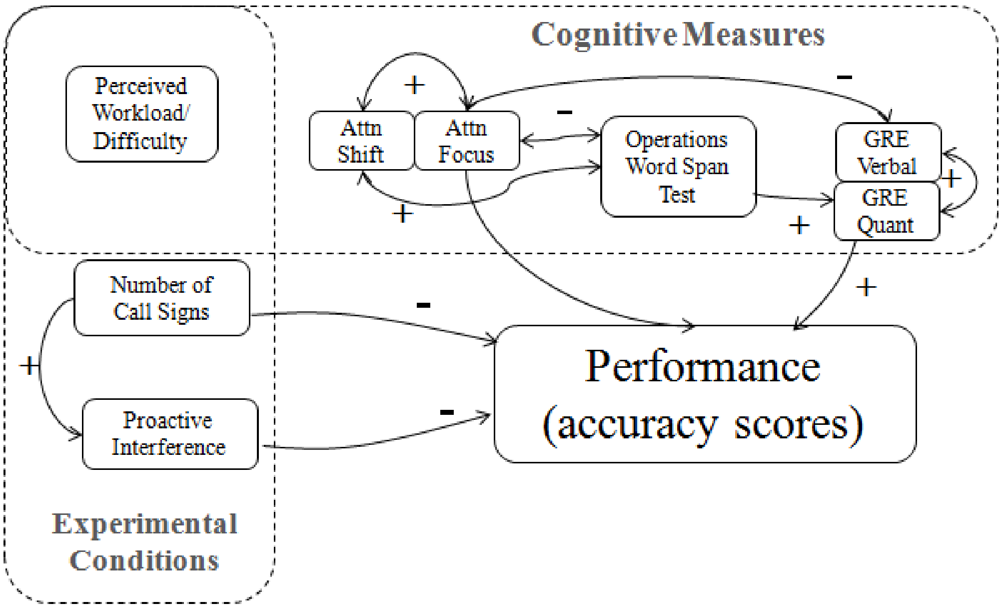

Figure 6. Causal diagram of experimental results.

Significance of Research

The results from this study indicate that proactive interference moderates the relationship between cognitive load and performance, meaning proactive interference has a stronger effect at higher cognitive loads. Considering the importance of this, the Air Force and its system designers should find ways to reduce proactive interference in multiaircraft communication tasks. Some ways to achieve this goal is to make call signs more distinct or temporally separate stimuli that could cause interference.

Given the current experimental constraints and participant's subjective opinions of their own performance, at a 95% accuracy rate, individuals are expected to be able to reliably monitor approximately six call signs. This operational capacity is subject to change as the task environment becomes more saturated and the operations tempo increases the cognitive demand of the operator. In a true operational setting, it is recommend that pilots operate two aircraft simultaneously then add more aircraft as competency is demonstrated.

The current study successfully bridged the gap between traditional working memory laboratory tests and real world applications by conducting a human subjects experiment on a multiaircraft control radio communications task. It found that these traditional cognitive assessments, specifically GRE scores, are predictive of accuracy on this aviation task. While results on the OSPAN test and Attention Control Scale are not significantly predictive of performance, at least the relationship is in the expected direction. However, more research should be conducted on this topic to find cognitive tests that are more predictive of performance on this task by extending an experiment similar to the one conducted in this research to include a broader sample of participants. This study laid the groundwork for a potential personnel selection tool where individuals' scores on cognitive ability assessments could be compared to their performance on certain measures in order to predict performance on a job like UAV operations. Other studies have been able to predict pilot training performance using measures of cognitive ability and personality traits (Carretta, et al., 2014), so further research should include personality tests, like the International Personality Item Pool (IPIP) which tests the Big 5 personality traits of openness, conscientiousness, extraversion, agreeableness, and neuroticism (Goldberg et al., 2006). The research by Rose et al. (2014) suggests that certain personality scores, like openness, could predict talent in RPA operations.

Recommendations for Future Research

While this research explored the basic cognitive factors associated with multiaircraft control, there are still several areas that could be further investigated. This study had a small sample size, so the first recommendation would be to conduct the experiment on more participants of varying education levels and backgrounds, resulting in a sample with a larger variance of scores. A larger sample size would allow relationships with moderate effect sizes to be statistically significant. Additionally, it would be useful to explore personality traits in relation to other variables tested in the study because they might be responsible for certain performance results. Since a ceiling effect was found in the accuracy score results, it would be useful to conduct a similar experiment where the participant is responsible for more call signs (until failure occurs) in order to achieve a more comprehensive view of how cognitive load affects performance.

Since the MMC software is able to log all actions the participant makes, it lends itself to several types of experiments. Users could manipulate the speed the radio calls play (at a constant or variable pace) and the number of call signs to create a workload model. The spatial audio component of the MMC system allows for voices to originate from up to 7 different locations in space. In a radio communications scenario, this could allow the user to place voices in certain spatial locations based on the particular mission. For Example, voice recognition software could automate radio calls for critical call signs to be placed in a specific spatial location different from distracter call signs (e.g. critical call signs in right ear, all other call signs in left ear).

Additionally, other cognitive tests like the Air Force Officer Qualifying Test (AFOQT), ACT, or SAT could be administered to participants to establish a tool that is predictive of performance on certain multiaircraft tasks. This could help organizations quickly select personnel who would be successful in a particular career field.

Summary

The current study investigated the cognitive aspects of multiaircraft control in a multiaircraft control radio communications task. It showed how much cognitive demand individuals can withstand while performing a task with interference present. The integration of these results could allow system designers to understand personal characteristics that predict an individual's performance on a multiaircraft control task and to design systems that minimize interference.

Appendix A

Instructions for Procedures

Attachment 2: Detailed Procedures

Before the Experiment

1. The participant and investigator sit down in the testing cubicle located in the human-systems integration laboratory (or a similar location in the laboratory at USAFA).
2. Investigator explains experimental purpose, goals, risks, and procedures to the participant.
   a. "Thank you for your participation in this study which is part of my AFIT Master's thesis. The goal of this experiment is to explore the cognitive aspects of multiaircraft control- meaning your ability to remember and respond when monitoring multiple UAVs. This project is mainly concerned with working memory, attention, and proactive interference (e.g., your ability to discard information in working memory so that it does not interfere with future tasks). In this experiment, you will be asked to complete four trials where you will listen to radio calls coming from the computer and you will have to respond to them by typing numbers in the keypad. You will also be asked to take a test designed to measure your working memory span. It involves solving a simple mathematical equation and memorizing a word. The entire experiment will take less than two hours from start to finish."
   b. "This experiment presents little risk. All data collected during the experiment and from the questionnaires will be kept private and no personally identifiable information will be recorded."
3. Informed consent
   a. "This is the informed consent document. It outlines the purpose of the study, risks, and procedures. You do not have to participate in this study and you may stop at any point during the study if you wish to not continue. Please read through it verbally indicate whether you would still like to participate in this study. If you have any questions, feel free to ask me."
   b. Participant reads the document and elects to proceed or to excuse themselves from the study.
4. Subjects are screened for vision (results will only be used for the purposes of this experiment.
   a. Simple eye chart to measure your visual acuity. "To test your vision, please hold the string up to the corner of your eye and read the letters in row 3, then 4, proceed until completed or you cannot proceed. You may wear your glasses or contacts for this as long as you wear them for the experimental trials."
      i. Record responses
   b. Ishihara color test "I will ask you to look at a series of images and you will indicate what shape you see in the dots."
      i. Record responses
   c. Please fill out questions 1-3 on the questionnaire.

   Experiment

5. Subjects shown how to use the MMC system on the laptop

a. "Right now, I will give you a quick 3 minute warm up to show you how to use the MMC system. Once the program starts, you will hear radio calls start playing. You will be given a list of call signs you are responsible for answering. Here are the call signs you will be listening for [show list of 3 call signs]. You may use the list to your left as a reference, but try to look only at the computer screen. If you hear radio calls for call signs that are not on the list, please type a zero into the chat window. If you are unsure whether you should respond to a call sign, type zero. An example radio call would be *"Ready Charlie go to blue one now." If you heard this and you were supposed to respond to it, you would look on the sheet, find "blue one" on the grid, and type in the number on the keyboard followed by enter[show grid and point to "blue one" and identify which number they would type]. You would continue to do this for every call sign you are responsible for. When the trial is done, you may take off your headphones.* Take a few seconds to memorize this list of words before we begin. When you are ready, put the headphones on and we will start.

b. Instruct subjects to wear the headphones with the wire on the left side of their head. "The headphones have been wiped down."

Operating directions for computer:

Post list of call signs to wall on right of the subject

1- start Openfire
2- open MMC
3- set radio frequency 100 at center position in MMC
4- right click frequency and select "add to workspace"
5- open chat box
6- Double click the Muc Logger icon to open the program
7- Open PlayStory and type in article number
8-start trial

c. Give a quick 3 minute warm-up scenario with 3 call signs (collect data)
d. Verify sound level

6. "Are you able to hear all of the radio calls at the volume it is now?"
7. Below is the experimental order as described in each trial:

Partici-pant Group		Short Trial	**Trial 1**	**Trial 2**	Break (15 minutes)	**Trial 3**	**Trial 4**
1	Call signs respon-sible for	3 mins	Charlie Gringo Laker Hopper Arrow ((11)) Tiger Eagle Barron Raptor	Laker Hopper Arrow Tiger Eagle ((12)) Charlie Gringo Barron Raptor	Operations word span test followed by a water/ restroom break	Charlie Gringo Laker Raptor Viking Arrow Tiger ((13)) Thunder Cobra	Charlie Gringo Laker Raptor Viking Thunder Cobra ((14)) Arrow Tiger

			Thunder Cobra Viking Shadow	Thunder Cobra Viking Shadow		Hopper Eagle Barron Shadow	Hopper Eagle Barron Shadow
2			Charlie Gringo Laker Raptor Viking Arrow Tiger ((13)) Thunder Cobra Hopper Eagle Barron Shadow	Charlie Gringo Laker Raptor Viking Thunder Cobra ((14)) Arrow Tiger Hopper Eagle Barron Shadow		Charlie Gringo Laker Hopper Arrow ((11)) Tiger Eagle Barron Raptor Thunder Cobra Viking Shadow	Laker Hopper Arrow Tiger Eagle ((12)) Charlie Gringo Barron Raptor Thunder Cobra Viking Shadow

8. Experimental Trial 1: Participants in odd subject number group will be given 5 call signs to which to attend and respond (e.g., Charlie, Ringo, Laker, Hopper, Arrow—**Article 11**) . Even numbered subjects will be given 7 call signs to which to attend and respond (e.g., Charlie, Ringo, Laker, Hopper, Arrow, Tiger, Eagle—**Article 13**).
   a. "You will now start the experimental trials. This trial will be approximately 8 minutes long. In the same way you did in the warm up, please listen for the call signs on this list [show list] and respond with the proper number. Take a few seconds to memorize this list of words before we begin. When you are ready, put the headphones on and we will start."
   b. Ask the subject to fill out the workload measurement question for trial 1 found in Attachment 3.
9. Experimental trial 2: Participants in odd subject number group will be given 5 call signs to which to attend and respond (e.g., Laker, Hopper, Arrow, Tiger, Eagle—**Article 12**). Even subject numbers will be given 7 call signs to which to listen to and respond (e.g., Ringo, Laker, Hopper, Arrow, Tiger, Eagle, Barron—**Article 14**).
   a. "This trial is similar to the last one, but now you will be responsible for these call signs [show list]. After this, you will be given a break and the working memory test. Take a few seconds to memorize this list of words before we begin. When you are ready, put the headphones on and we will start."
   b. Ask the subject to fill out the workload measurement question for trial 2 found in Attachment 3.
10. After completing the first two trials, participants will be given a break. During the break, an operations word span test is administered to subjects on the computer
   a. The description of the operations word span test is explained to participants
      i. "Right now, you will be taking the operations word span test which is designed to measure your working memory capacity. Although this task will sound easy, many people find it difficult. Just do your best. You will view a series of PowerPoint slides

that contain a mathematical operation followed by word on the same slide. Your job is to read the mathematical expression aloud and say whether it is correct or incorrect. Additionally, you must remember the word that follows it. An example of what you should say is ‘four divided by two plus five equals six, incorrect.’ Then at the end of the trial, you will see a question mark on the screen. This indicates that you should say all of the words you were asked to remember in that trial. You do not have to say the words in the order which they were presented. After reciting all of the words, you will move on to the next trial. You will first be asked to remember 2 words and will work your way up to 6 words. You will do two trials at each word level, so for example, you will be presented with three math expression/word combinations twice before you move on to four words. Next, you will be given you a quick, two word trial so that you can see how this works. Do you have questions about this so far?” (answer any questions the participant has)

b. Participants complete a 2 word trial (slides 1-5)
c. Investigator gives participant feedback on their performance and answers any questions they may have
d. Experimental trials start. All data about “correct/incorrect” answers and words will be recorded on a score sheet (Attachment 6) during trials by the investigator.
   i. Give subjects two word trial #1 and #2
   ii. Give subjects three word trial #1 and #2
   iii. Give subjects four word trial #1 and #2
   iv. Give subjects five word trial #1 and #2
   v. Give subjects six word trial #1 and #2

11. Experimental trial 3: Participants in odd subject number group will be given 7 call signs (Ringo, Laker, Hopper, Arrow, Tiger, Eagle, Barron—**Article 13**). Even subject numbers will be given 5 call signs (Laker, Hopper, Arrow, Tiger, Eagle—**Article 11**).
    a. “This trial is similar to the others, but now you will be responsible for these call signs [show list]. Take a few seconds to memorize this list of words before we begin. When you are ready, put the headphones on and we will start.”
    b. Ask the subject to fill out the workload measurement question for trial 1 found in Attachment 3.
12. Experimental trial 4: Participants in odd subject number group will be given 7 call signs (e.g., Charlie, Ringo, Laker, Hopper, Arrow, Tiger, Eagle—**Article 13**). Even subject numbers will be given 5 call signs(e.g., Charlie, Ringo, Laker, Hopper, Arrow—**Article 11**) (Laker, Hopper, Arrow, Tiger, Eagle).
    a. “This is the final trial. Again, this trial is similar to the others, but now you will be responsible for these call signs [show list]. Take a few seconds to memorize this list of words before we begin. When you are ready, put the headphones on and we will start.”
    b. Ask the subject to fill out the workload measurement question for trial 1 found in Attachment 3.
13. Administer rest of questionnaire (Attachment 3)
14. Subject debriefed and dismissed.

a. “Thank you so much for participating in this study. I appreciate your help. If you have any more questions about the study, you can email me at Kelly.amaddio@afit.edu or Dr. Michael Miller at michael.miller@afit.edu or come back to this lab.

Appendix B

Questionnaire

Before the experiment:

Vision: ____

Color: _____

1) Are you fluent in English?
   Yes_____ No _____

2) Please indicate your age: _______years
3) Please indicate your gender: _____ Male _____ Female

After each experimental trial:

Trial 1: Circle ONE rating that best indicates your workload for the just-completed trial.

1. Under-utilized: Nothing to do. Rather boring.
2. Relaxed: More than enough time for all the tasks. Active on the task.
3. Comfortably busy pace: All tasks well in hand. Busy but stimulated. Could keep going continuously at this level.
4. High: Non-essential tasks suffering. Could not work at this level.
5. Excessive: Behind on tasks. Losing track of full picture.

Trial 1: Circle ONE rating that best indicates the difficulty of the just-completed trial.

1 Very easy	**2** Easy	**3** Neutral	**4** Difficult	**5** Very difficult

Trial 2: Circle ONE rating that best indicates your workload for the just-completed trial.

1. Under-utilized: Nothing to do. Rather boring.
2. Relaxed: More than enough time for all the tasks. Active on the task.
3. Comfortably busy pace: All tasks well in hand. Busy but stimulated. Could keep going continuously at this level.
4. High: Non-essential tasks suffering. Could not work at this level.
5. Excessive: Behind on tasks. Losing track of full picture.

Trial 2: Circle ONE rating that best indicates the difficulty of the just-completed trial.

1 Very easy	**2** Easy	**3** Neutral	**4** Difficult	**5** Very difficult

Trial 3: Circle ONE rating that best indicates your workload for the just-completed trial.

1. Under-utilized: Nothing to do. Rather boring.
2. Relaxed: More than enough time for all the tasks. Active on the task.
3. Comfortably busy pace: All tasks well in hand. Busy but stimulated. Could keep going continuously at this level.
4. High: Non-essential tasks suffering. Could not work at this level.
5. Excessive: Behind on tasks. Losing track of full picture.

Trial 3: Circle ONE rating that best indicates the difficulty of the just-completed trial.

1 Very easy	**2** Easy	**3** Neutral	**4** Difficult	**5** Very difficult

Trial 4: Circle ONE rating that best indicates your workload for the just-completed trial.

1. Under-utilized: Nothing to do. Rather boring.
2. Relaxed: More than enough time for all the tasks. Active on the task.
3. Comfortably busy pace: All tasks well in hand. Busy but stimulated. Could keep going continuously at this level.
4. High: Non-essential tasks suffering. Could not work at this level.
5. Excessive: Behind on tasks. Losing track of full picture.

Trial 4: Circle ONE rating that best indicates the difficulty of the just-completed trial.

1 Very easy	**2** Easy	**3** Neutral	**4** Difficult	**5** Very difficult

Based on your experience today, how many call signs do you think you could monitor comfortably before you would begin missing time critical information? ______

Attentional Control Questionnaire:

Read each statement carefully and decide how well it describes you. For each statement, respond by circling the response that best represents you using the following choices:

1 Almost never	**2** sometimes	**3** often	**4** always

1. It's very hard for me to concentrate on a difficult task when there are noises around.

1 Almost never	**2** sometimes	**3** often	**4** always

2. When I need to concentrate and solve a problem, I have trouble focusing my attention.

1 Almost never	**2** sometimes	**3** often	**4** always

3. When I am working hard on something, I still get distracted by events around me.

1 Almost never	**2** sometimes	**3** often	**4** always

4. My concentration is good even in there is music in the room around me.

1 Almost never	**2** sometimes	**3** often	**4** always

5. When concentrating, I can focus my attention so that I become unaware of what's going on in the room around me.

1 Almost never	**2** sometimes	**3** often	**4** always

6. When I am reading or studying, I am easily distracted if there are people talking in the same room.

1 Almost never	**2** sometimes	**3** often	**4** always

7. When trying to focus my attention on something, I have difficulty blocking out distracting thoughts.

1 Almost never	**2** sometimes	**3** often	**4** always

8. I have a hard time concentrating when I'm excited about something.

1 Almost never	**2** sometimes	**3** often	**4** always

9. When concentrating, I ignore feelings of hunger or thirst.

1 Almost never	2 sometimes	3 often	4 always

10. I can quickly switch from one task to another.

1 Almost never	2 sometimes	3 often	4 always

11. It takes me a while to get really involved in a new task.

1 Almost never	2 sometimes	3 often	4 always

12. It is difficult for me to coordinate my attention between the listening and writing required when taking notes during lectures.

1 Almost never	2 sometimes	3 often	4 always

13. I can become interested in a new topic very quickly when I need to.

1 Almost never	2 sometimes	3 often	4 always

14. It is easy for me to read or write while I'm also talking on the phone.

1 Almost never	2 sometimes	3 often	4 always

15. I have trouble carrying on two conversations at once.

1 Almost never	2 sometimes	3 often	4 always

16. I have a hard time coming up with new ideas quickly.

1 Almost never	2 sometimes	3 often	4 always

17. After being interrupted or distracted, I can easily shift my attention back to what I was doing.

1 Almost never	2 sometimes	3 often	4 always

18. When a distracting thought comes to mind, it is easy for me to shift my attention away from it.

1 Almost never	2 sometimes	3 often	4 always

19. It is easy for me to alternate between two different tasks.

1 Almost never	2 sometimes	3 often	4 always

20. It is hard for me to break away from one way of thinking about something and look at it from another point of view.

1 Almost never	2 sometimes	3 often	4 always

Appendix C

OSPAN Score Sheet

Trial	Mathematical expression #1	Correct/ incorrect?	Right/ wrong?	Word	Score	Mathematical expression #2	Correct/ incorrect?	Right/ wrong?	Word	Score
1	(9/3) + 4 = 7	correct		garage		(3 x 3) - 6 = 3	correct		ketchup	
	(3 x 2) + 1 = 7	correct		monkey		(2/1) + 2 = 4	correct		tattoo	
2	(6/3) + 3 = 9	incorrect		forest		(8/2) + 3 = 4	incorrect		jacket	
	(3 x 1) + 4 = 7	correct		rocket		(7 x 1) - 4 = 9	incorrect		mother	
	(2/1) + 2 = 4	correct		scissors		(8/4) + 3 = 5	correct		pasta	
3	(4/2) + 7 = 3	incorrect		pizza		(4 x 2) - 4 = 4	correct		costume	
	(3 x 2) - 4 = 9	incorrect		yellow		(9/3) + 4 = 7	correct		hotel	
	(5/1) + 2 = 7	correct		pretzel		(2 x 3) - 1 = 9	incorrect		tunnel	
	(4/2) + 7 = 3	incorrect		wedding		(3 x 3) - 4 = 9	incorrect		baby	
4	(3 x 2) - 4 = 9	incorrect		blizzard		(8/1) - 5 = 4	incorrect		kitchen	
	(3 x 3) - 6 = 3	correct		garden		(6/2) + 6 = 9	correct		sausage	
	(2/1) + 2 = 4	correct		burger		(4 x 2) - 3 = 9	incorrect		pumpkin	
	(8/2) + 3 = 4	incorrect		money		(1 x 2) + 7 = 9	correct		castle	
	(7 x 1) - 4 = 9	incorrect		shadow		(3 x 2) - 4 = 9	incorrect		lemon	
5	(8/4) + 3 = 5	correct		sister		(5/1) + 2 = 7	correct		luggage	
	(4 x 2) - 4 = 4	correct		tiger		(4/2) + 7 = 3	incorrect		vacuum	
	(9/3) + 4 = 7	correct		carrot		(3 x 2) - 4 = 9	incorrect		picture	
	(2 x 3) - 1 = 9	incorrect		printer		(9/3) + 4 = 7	correct		flower	
	(3 x 3) - 4 = 9	incorrect		menu		(3 x 2) + 1 = 7	correct		hockey	
	(8/1) - 5 = 4	incorrect		hamster		(6/3) + 3 = 9	incorrect		pudding	
6	(6/2) + 6 = 9	correct		color		(3 x 1) + 4 = 7	correct		rabbit	
	(4 x 2) - 3 = 9	incorrect		wallet		(2/1) + 2 = 4	correct		navy	
	(1 x 2) + 7 = 9	correct		marker		(4/2) + 7 = 3	incorrect		cousin	
	(3 x 2) - 4 = 9	incorrect		sweater		(3 x 2) - 4 = 9	incorrect		silver	
	(5/1) + 2 = 7	correct		kitten		(5/1) + 2 = 7	correct		Candle	
	(4/2) + 7 = 3	incorrect		army		(4/2) + 7 = 3	incorrect		Angel	
	(3 x 2) - 4 = 9	incorrect		tennis		(3 x 2) - 3 = 9	incorrect		pepper	

Appendix D

Consent Document

Greetings! You are being asked to take part in a research study carried out by 2Lt Kelly Amaddio, AFIT Masters Student in Engineering Management. This form explains the study and your part in it if you decide to participate. Please read the form carefully; take as much time as desired. Ask the researcher to explain anything you do not understand. You can decide not to participate in the study. If you participate in the study, you can change your mind later or quit at any time, without any penalty or loss of services or benefits.

Study Title: The Cognition of Multi-Aircraft Control (MAC): Proactive Interference and Working Memory Capacity Impact on Voice Communication in MAC

Primary Researcher:

Name	Title/Department	E-mail	Telephone
Kelly Amaddio	Masters Student in Engineering Management, AFIT/ENV	Kelly.Amaddio@afit.edu	937.255.3636, x4730

What is this study about? The purpose of this study is to determine how many call signs a UAV pilot can attend to by evaluating performance during a multi-aircraft radio monitoring task. These results will inform system designers and policy makers regarding the number of aircraft an individual pilot can control given current technology limitations. The entire study should take less than 2 hours.

What will I be asked to do if I participate in this study? If you take part in the study, you will complete a short hearing and vision survey. You will then participate in a short baseline trial lasting approximately 5 minutes, four 10-minute radio monitoring task trials, a 15-min working memory test (Operations word span test), and a post-experimental survey that will take less than 15 minutes to complete. During all experimental trials, an eye tracker may be used to measure your pupil diameter, eye movements, and blink rate. This apparatus will be non-invasive and will not keep any recordings of your eyes. The entire experiment will take place in the Human Systems integration Lab (where you are now). For the radio monitoring tasks, you will be given a list of call signs that you will be responsible for monitoring. You will hear radio calls on the headphones and will have to respond to your call signs by typing a number you find on a grid. The working memory test will occur between the second and third trials of the radio monitoring task. During this test, you will be asked to verify the correctness of a simple mathematical problem while remembering a set of words.

Are there any benefits to me if I participate in this study? The main benefits of this study will be to help provide a foundation for informing system designers and policy makers regarding the number of aircraft an individual pilot can control given current technology limitations.

Are there any risks to me if I participate in this study? Because the working memory test requires you to complete simple math and recall a group of words, it might become frustrating to you. Therefore, you may stop at any time.

Will my information be kept private? The data for this study will be kept confidential to the extent allowed by federal and state law. We will not record any personally identifiable information and your name will not be associated with the findings. The digital file containing the survey and data collection results, as well as the study write-up will be secured on a password-protected computer assigned to the researcher. For the sake or organization, you will be assigned a subject number which will not be connected to your name.

Your information will only be released, if requested, to authorized members of the AFIT Institutional Review Board, to ensure research compliance with federal and state law. Your information will not be released to any other entity. The results of this study may be published or presented at professional meetings, but the identities of all research participants will not be collected and will therefore remain anonymous. The data for this study will be kept for three years, as required by AFIT policy, after which time the digital file containing the data will be destroyed.

Are there any costs or payments for participating in this study? There will be no costs or payments to you for taking part in this study.

Who can I talk to if I have questions? If you have questions about this study or the information in this form, please contact the researcher using the contact information provided above. If you have questions about your rights as a research participant, or would like to report a concern or complaint about this study, please contact the WPAFB Institutional Review Board at (937) 255-3636, x4730 or e-mail HumanSubjects@afit.edu, or regular mail at: Wright Research Site IRB, 711 HPW/IR, 2245 Monahan Way, Wright-Patterson AFB, OH 45433

What are my rights as a research study volunteer? Your participation in this research study is completely voluntary. You may choose not to be a part of this study. There will be no penalty to you if you choose not to participate. You may choose not to answer specific questions or to stop participating at any time.

Appendix E

Curriculum Vita

Kelly M. Amaddio, 2nd Lt, USAF

CONTACT INFORMATION

Address: Department of Systems Engineering and Management
Air Force Institute of Technology
2950 Hobson Way
WPAFB, OH 45433-7765

Phone: 502-718-3459

E-mail Address: Kelly.amaddio@afit.edu

MILITARY INFORMATION

Clearance:	Secret (July 2009)
Rank:	Second Lieutenant
Date of Rank:	29 May 2013
Primary AFSC:	61B
Duty AFSC:	9FEF

FORMAL EDUCATION

M.S. Engineering Management, Human Factors Concentration, 2015 (*in progress*)
Air Force Institute of Technology, Wright-Patterson AFB, OH
Grade Point Average: 3.86 of 4.0
Thesis: *The Cognition of Multi-Aircraft Control (MAC)*

B.S. Behavioral Sciences, 2013
United States Air Force Academy, Colorado Springs, CO
Grade Point Average: 3.2 of 4.0

POSTER PRESENTATIONS

Amaddio, K., Miller, M., (2014). "The Cognition of Multi-Aircraft Control (MAC): Proactive Interference and Working Memory Capacity Impact on Voice Communication in MAC," AFRL-AFIT Colloquium - Human-Machine Systems 2.0, Wright-Patterson, OH, Sep 30, 2014.

PROFESSIONAL AND HONOR SOCIETIES

2012 - Present *Psi Chi*, Psychology Honors Society, Vice President
2014 - Present *Sigma Iota Epsilon*, Professional Management Fraternity, Treasurer
2004 - Present *Tau Beta Pi*, International Honor Society for Engineers

Appendix F

22 October 2014

MEMORANDUM FOR AFIT IRB REVIEWER

FROM: Dr. Michael E. Miller; AFIT/ENV

SUBJECT: Request for exemption from human experimentation requirements (32 CFR 219, DoDD 3216.2 and AFI 40-402) for *The Cognition of Multi-Aircraft Control (MAC): Proactive Interference and Working Memory Capacity Impact on Voice Communication in MAC*.

1. The purpose of this study is to determine how many call signs a UAV pilot can attend to by evaluating performance during a multi-aircraft radio monitoring task. These results will inform system designers and policy makers regarding the number of aircraft an individual pilot can control given current technology limitations. The results will be published as a master's thesis.

2. This request is based on the Code of Federal Regulations, title 32, part 219, section 101,
paragraph (b) (2) Research activities that involve the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures, or observation of public behavior unless: (i) Information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; and (ii) Any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, or reputation. The following information is provided to show cause for such an exemption:

a) Equipment and facilities: Experiment will be conducted in the Air Force Institute of Technology Human Systems Integration laboratory and/or at the Air Force Academy's Behavioral Science Research Laboratory. The Mirametrix S2 Eye Tracker will be used to remotely measure pupil diameter, eye movements, and blink rate. No recordings of the likeness of the participants can or will be retained from this apparatus; the only output will be a Matlab data file containing eye movements and pupil diameter. A computer will further be used to display text prompts to support a working memory test and audio prompts simulating radio traffic. Further, a computer will be used to display a grid of potential responses and to record the participants' responses to relevant call signs. Reaction times and the participants responses to queries based on audible or visually presented stimuli will be recorded.

b) Subjects: 20 to 40 male and female volunteers age 18-45 with 20/20 eyesight (or corrected to 20/20 vision), and no color blindness will participate in this study. There is no educational requirement, although most participants will be undergraduate or graduate college students, due to availability of subjects. Once the exemption is accepted, a recruitment e-mail message will be sent to AFIT students and other officers on base.

c) Timeframe: The study will take place over the course of approximately six months between October 2014 and March 2015.

d) Data collected: Demographic data such as age, gender, and education level will be collected from each participant. The participants will also take part in color blindness and visual acuity measures. The results from these tests will not be shared with the participants and they will complete the experiment regardless of the outcome of the tests. During the radio call trials, data about radio call responses will be collected. During the break, a working memory test (called the Operations Word Span test) will be administered. A video-based eye tracker will be used to track the user's eye movements to both determine where they are looking and to monitor their pupil size during the working memory test and the experimental trials. This data will be examined in relation to the subjects' performance and subjective workload. This eye tracker records data about the eye location and pupil size of the participant but does not permit recording of the user's face. Additionally, a questionnaire will be used to capture each subject's opinions about the experiment's workload level, and other subjective impressions once the experimental trials are complete. (Attachment 1). No personally identifiable information will be collected.

e) Risks to Subjects: The participants will not be exposed to any risk beyond those experienced within their every day working environment.

f) Informed consent: All subjects are self-selected to volunteer to participate in the study. No adverse action is taken against those who choose not to participate or to terminate participation prior to study completion. Subjects are made aware of the nature and purpose of the research and sponsors of the research. A copy of the Privacy Act Statement of 1974 is presented for their review.

4. I understand that I will not collect names or any personally identifiable information from each participant. The data collected will be protected at all times, only be known to the researchers, and managed according to the AFIT interview protocol. All interview data will only be handled by Lieutenant Kelly Amaddio and Dr. Michael E. Miller. At the conclusion of the study, all data will be turned over to Dr. Michael Miller and all other copies will be destroyed. Also, if a subject's future response reasonably places them at risk of criminal or civil liability or is damaging to their financial standing, employability, or reputation, I understand that I am required to immediately file an adverse event report with the IRB office.

5. If you have any questions about this request, please contact Dr. Michael E. Miller (primary investigator) – Phone 937-255-3636, ext. 4651; E-mail – michael.miller@afit.edu.

Dr. Michael E. Miller
Principal Investigator

Bibliography

About the GRE® revised General Test. (n.d.). Retrieved February 5, 2015, from http://www.ets.org/gre/revised_general/about?WT.ac=grehome_greabout_b_130807.

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science, 15(2),* 106-111.

Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(3), 570.

Baddeley, A. (2000). The episodic buffer: a new component of working memory?. *Trends in Cognitive Sciences, 4(11)*, 417-423.

Baddeley, A.D. and Hitch, G. (1974) *Working memory*. In The Psychology of Learning and Motivation (Bower, G.A., ed.), pp. 48-79, Academic Press.

Blair, E. A., Rahill, K. M., Finomore, V., Satterfield, K., Shaw, T., & Funke, G. (2014, September). Best of Both Worlds Evaluation of Multi-Modal Communication Management Suite. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 58, No. 1, pp. 410-414). SAGE Publications.

Carretta, T. R., Teachout, M. S., Ree, M. J., Barto, E. L., King, R. E., & Michaels, C. F. (2014). Consistency of the Relations of Cognitive Ability and Personality Traits to Pilot Training Performance. *The International Journal of Aviation Psychology*, *24*(4), 247-264.

Cohen, J. (1992). A power primer. *Psychological bulletin*, *112*(1), 155.

Colombi, J., Miller, M., Schneider, M., McGrogan, J., Long, D., & Plaga, J. (2012). Predictive mental workload modeling for semi-autonomous system design: Implications for systems of systems. *Systems Engineering, 15*(4), 448-460.

Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic bulletin & review*, *12*(5), 769-786.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450-466.

Derryberry, D., & Reed, M. A. (2002). Anxiety-related attentional biases and their regulation by attentional control. *Journal of abnormal psychology*, *111*(2), 225.

Engle, R. W., & Oransky, N. (1999). Multi-store versus dynamic models of temporary storage in memory. *The Nature of Cognition*, 515-555.

ETS (2014). *Performance statistics on the GRE revised general test.* Retrieved February 10, 2015 from https://www.ets.org/s/gre/pdf/gre_guide_table1a.pdf.

Fajkowska, M., & Derryberry, D. (2010). Psychometric properties of Attentional Control Scale: The preliminary study on a Polish sample. *Polish Psychological Bulletin*, *41*(1), 1-7.

Finomore, V.S., Popik, D.K., Castle, C.E., & Dallman, R.C. (2010). Effects of network-centric multi-modal communication on a communication-monitoring task. *Proceedings of the Human Factors and Ergonomics Society, 54,* 2125- 2129.

Finomore, V., Stewart, J., Singh, R., Raj, B., & Dallman, R. (2012). Demonstration of advanced multi-modal, network-centric communication management suite. *Proceedings of the Interspeech, 13.*

Finomore, V., Satterfield, K., Sitz, A., Castle, C., Funke, G., Shaw, T., & Funke, M. (2012). Effects of the multi-modal communication tool on communication and change detection for command and control operators. *Proceedings of the Human Factors and Ergonomics Society, 56, 1464-1465*.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. C. (2006). The International Personality Item Pool and the future of public-domain personality measures. *Journal of Research in Personality, 40*, 84-96.

Handwerk, B. (2013, December 2). 5 Surprising Drone Uses (Besides Amazon Delivery). Retrieved February 12, 2015, from http://news.nationalgeographic.com/news/2013/12/131202-drone-uav-uas-amazon-octocopter-bezos-science-aircraft-unmanned-robot/.

Hofmann, D. A. (1997). An overview of the logic and rationale of hierarchical linear models. *Journal of management*, *23*(6), 723-744.

Ishihara, S. (1980). *Ishihara's design charts for colour-blindness of unlettered persons*. Kanehara & Company.

Kane, M. J., Bleckley, M. K., Conway, A. R., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General, 130(2)*, 169.

Kane, M. J., & Engle, R. W. (2000). Working-memory capacity, proactive interference, and divided attention: limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26(2)*, 336.

Kumar, S., Joseph, S., Pearson, B., Teki, S., Fox, Z. V., Griffiths, T. D., & Husain, M. (2013). Resource allocation and prioritization in auditory working memory. *Cognitive Neuroscience, 4(1)*, 12-20.

Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2001). A comprehensive meta-analysis of the predictive validity of the graduate record examinations: implications for graduate student selection and performance. *Psychological bulletin*, *127*(1), 162.

Logarithmic Near Visual Acuity Chart 2000 "New ETDRS" Charts (2011). Precision Vision.

May, C. P., Hasher, L., & Kane, M. J. (1999). The role of interference in memory span. *Memory & Cognition, 27(5)*, 759-767.

McGrogan, J., Schneider, M., Wirthlin, R., Columbi, J., & Miller, M. (2011). Predicting the impact of multi-aircraft control on UAS operations. In T. Doolen, B. Editor, & E. Van Aken. *Proceedings of the 2011 Industrial Engineering Research Conference*.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review, 63(2),* 81.

New York Times (2013), "Obama's Speech on Drone Policy," retrieved from http://www.nytimes.com/2013/05/24/us/politics/transcript-of-obamas-speech-on-drone-policy.html?pagewanted=all&_r=0.

Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience, 139(1),* 23-38.

Rose, M. R., Barron, L. G., Carretta, T. R., Arnold, R. D., & Howse, W. R. (2014). Early Identification of Unmanned Aircraft Pilots Using Measures of Personality and Aptitude. *The International Journal of Aviation Psychology*, *24*(1), 36-52.

Sweller, J (June 1988). "Cognitive load during problem solving: Effects on learning". *Cognitive Science 12 (2)*: 257–285.

Tattersall, A. J., & Foord, P. S. (1996). An experimental evaluation of instantaneous self-assessment as a measure of workload. *Ergonomics*, *39*(5), 740-748.

Teigen, K. H. (1994). Yerkes-Dodson: A law for all seasons. Theory & Psychology, 4(4), 525-547.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of memory and language*, *28*(2), 127-154.

Unmanned Systems Integrated Roadmap FY2013-2038. United States Department of Defense. Reference Number: 14-S-0553.

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta psychologica*, *41*(1), 67-85.

Wickens, C. D. (1981). *Processing Resources in Attention, Dual Task Performance, and Workload Assessment* (No. EPL-81-3/ONR-81-3). Illinois Univ at Urbana Engineering-Psychology Research Lab.

Woltman, H., Feldstain, A., MacKay, J. C., & Rocchi, M. (2012). An introduction to hierarchical linear modeling. *Tutorials in Quantitative Methods for Psychology*, *8*(1), 52-69.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From – To)*
26 Mar 2015	Master's Thesis	04 Aug 2013 - 26 Mar 2015

4. TITLE AND SUBTITLE
The Cognition of Multiaircraft Control (MAC): Cognitive Ability Predictors, Working Memory, Interference, and Attention Control in Radio Communication

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Kelly M. Amaddio, Second Lieutenant, USAF

5d. PROJECT NUMBER
N/A

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)
Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/ENV)
2950 Hobson Way, Building 640
WPAFB OH 45433-8865

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT-ENV-MS-15-M-205

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Col Anthony Tvaryanas, MD
711th Human Performance Wing/HP SG
2510 Fifth St Building 840 room W413.32
Wright-Patterson AFB, OH 45433
Sponsor's telephone (937) 798-3253, and
anthony.tvaryanas@us.af.mil

10. SPONSOR/MONITOR'S ACRONYM(S)
711th HPW/HP SG

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
DISTRIBUTION STATEMENT A. APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
As the number of U.S. Air Force missions requiring UAVs has rapidly increased without commensurate increases in manpower, systems which permit a single operator to supervise and control multiple, highly-automated aircraft are being considered. The operator of such a system may be required to monitor and respond to voice communications for multiple UAVs, each of which can have aircraft specific call signs. The need to monitor this array of call signs may impose excessive requirements on constrained operator attention, working memory, and cognitive processing. The current research investigates the cognitive load (number of aircraft call signs) an individual can handle and explores the effect of proactive interference (PI) within this application. The results indicate a reduction in performance as the number of call signs are increased from 5 to 7 in the presence of PI. Additionally, this study seeks to understand if individual differences in working memory and attention predict performance on the multiaircraft control radio communication task through the application of the Operations Word Span test, Attention Control Scale, and GRE scores. Hierarchical linear modeling was used to determine the relationships among these and other variables.

15. SUBJECT TERMS
Multiaircraft Control, Attention control, Working memory, Proactive interference

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Researcher, Michael E. Miller, Ph.D, USAF
U	U	U	UU or SAR	88 Pages	**19b. TELEPHONE NUMBER** *(Include area code)* (937) 255-3636, x 4730 (kelly.amaddio@afit.edu)